Instytut Informatyki Teoretycznej i Stosowanej
Polskiej Akademii Nauk

Rozróżnialność pomiarów kwantowych

Rozprawa doktorska

mgr Aleksandra Krawiec
Promotor:
dr hab. Zbigniew Puchała, prof. IITiS
Promotor pomocniczy:
dr hab. inż. Łukasz Pawela, prof. IITiS

Gliwice, 17.02.2023

Institute of Theoretical and Applied Informatics, Polish Academy of Sciences

Discrimination of quantum measurements

Doctoral dissertation

mgr Aleksandra Krawiec
Supervisor:
dr hab. Zbigniew Puchała, prof. IITiS
Co-supervisor:
dr hab. inż. Łukasz Pawela, prof. IITiS

Gliwice, 17.02.2023

Contents

List of publications

Publications relevant to this dissertation are highlighted bold.

1. A. Krawiec, Ł. Pawela, Z. Puchała; *Discrimination and certification of unknown quantum measurements*; arXiv preprint arXiv:2301.04948
2. A. Glos, A. Krawiec, Z. Zimborás; *Space-efficient binary optimization for variational quantum computing*; npj Quantum Information, vol. 8, issue 1, 2022
3. **A. Krawiec, Ł. Pawela, Z. Puchała; *Excluding false negative error in certification of quantum channels*; Scientific Reports, vol. 11, pp. 21716, 2021**
4. **Z. Puchała, Ł. Pawela, A. Krawiec, R. Kukulski, M. Oszmaniec; *Multiple-shot and unambiguous discrimination of von Neumann measurements*; Quantum, vol. 5, p. 425, 2021**
5. A. Glos, A. Krawiec, Ł. Pawela; *Asymptotic entropy of the Gibbs state of complex networks*; Scientific Reports, vol. 11, no.1, pp. 1-9, 2021
6. **P. Lewandowska, A. Krawiec, R. Kukulski, Ł. Pawela, Z. Puchała; *On the optimal certification of von Neumann measurements*; Scientific Reports, vol. 19, no.1, pp. 1-16, 2021**
7. **A. Krawiec, Ł. Pawela, Z. Puchała; *Discrimination of POVMs with rank-one effects*; Quantum Information Processing, vol. 19, 2020**
8. **Z. Puchała, Ł. Pawela, A. Krawiec, R. Kukulski; *Strategies for optimal single-shot discrimination of quantum measurements*; Physical Review A, vol. 98, issue 4, 2018**
9. A. Glos, A. Krawiec, R. Kukulski, Z. Puchała; *Vertices cannot be hidden from quantum spatial search for almost all random graphs*; Quantum Information Processing, vol. 17, pp. 81, 2018
10. Z. Puchała, Ł. Rudnicki, A. Krawiec, K. Życzkowski; *Majorization uncertainty relations for mixed quantum states*; Journal of Physics A: Mathematical and Theoretical, vol. 51, issue 17, 2018

Abstract in Polish

W tej rozprawie wykazujemy, że adaptacyjne schematy rozróżniania mogą poprawić rozróżnialność pomiarów kwantowych. Istnieją trzy podstawowe podejścia do badania problemu rozróżnialności: rozróżnianie symetryczne, jednoznaczne i asymetryczne.

Celem rozróżniania symetrycznego jest zminimalizowanie prawdopodobieństwa podjęcia błędnej decyzji, a więc maksymalizacja prawdopodobieństwa, że rozróżnianie się powiodło. Drugie podejście nazywane jest rozróżnieniem jednoznacznym. Wykorzystując to podejście, jeżeli otrzymamy rozstrzygający wynik rozróżniania, możemy być pewni, że jest on poprawny. Istnieje jednak szansa uzyskania wyniku nierozstrzygającego. Trzeci podejście, jakim jest rozróżnianie asymetryczne, znane jest również jako certyfikacja i opiera się na statystycznym testowaniu hipotez. W tym podejściu rozważamy oddzielnie błędy fałszywie dodatnie i fałszywie ujemne.

W podstawowym schemacie rozróżniania zakładamy, że jeden z dwóch pomiarów kwantowych, których opisy klasyczne są nam znane, jest potajemnie wybrany i schowany w czarnej skrzynce. Tej skrzynki nie można otworzyć, ale możliwe jest użycie pomiaru, który znajduje się wewnątrz niej. Przygotowujemy jako stan wejściowy stan kwantowy, który może być splątany z dodatkowym systemem. Następnie jest on mierzony przy pomocy pomiaru znajdującego się w czarnej skrzynce. Na podstawie otrzymanej etykiety pomiaru dokonujemy pomiaru dodatkowego systemu. Jego wynik pozwala na podjęcie decyzji, który z pomiarów znajdował się w czarnej skrzynce.

Wszystkie trzy podejścia do problemu rozróżniania są w pierwszej kolejności badane w sytuacji, gdy czarna skrzynka, która zawiera jeden z dwóch pomiarów, może być dostępna tylko raz. Potem badamy również równoległy schemat rozróżniania oraz najbardziej ogólny schemat adaptacyjny, który dopuszcza wykonywanie dodatkowych procedur pomiędzy kolejnymi zapytaniami do czarnej skrzynki. Dzięki temu możliwa jest modyfikacja stanu wejściowego dla kolejnego zapytania. Teza tej rozprawy brzmi: *Schematy adaptacyjne mogą ulepszyć rozróżnianie pomiarów kwantowych.*

Niniejsza rozprawa składa się z siedmiu rozdziałów i dwóch dodatków. Rozdział 1 zawiera wprowadzenie i motywację, jaka towarzyszyła napisaniu tej pracy.

Preliminaria matematyczne, podstawowe pojęcia kwantowej teorii informacji i wprowadzenie narzędzi matematycznych znajdują się w rozdziale 2. Rozróżnianie symetryczne pomiarów kwantowych jest badane w sytuacjach pojedynczych i wielokrotnych odpowiednio w rozdziałach 3 i 4. Badania rozróżnienia jednoznacznego znajdują się w rozdziale 5. Rozdział 6 jest poświęcony rozróżnianiu asymetrycznemu. Końcowe uwagi i wnioski można znaleźć w rozdziale 7. Na końcu pracy znajdują się dwa dodatki, które zawierają dowody twierdzeń, które nie znalazły się w tekście głównym.

Abstract in English

In this dissertation, we demonstrate that adaptive discrimination schemes can improve the discrimination of quantum measurements. The problem of discrimination of quantum measurements is studied in three approaches, which are: symmetric, unambiguous and asymmetric discrimination.

Symmetric discrimination is also known as minimum error discrimination. Its goal is to minimize the probability of making an erroneous decision, thus maximizing the probability that the discrimination is correct. The second approach is called unambiguous discrimination. In this approach, whenever we get the conclusive result of discrimination, we can be sure that it is correct. However, there is a chance of getting an inconclusive result of discrimination. The third approach, which is asymmetric discrimination, is also known as certification, and it is based on statistical hypothesis testing. This time, we consider the false positive and false negative errors separately.

In the basic discrimination scheme, we assume that one of two quantum measurements, which both classical descriptions are known, is secretly chosen and hidden in a black box. This box cannot be opened, but we can use measurement inside the black box. We prepare as input some quantum state, which can be entangled with some additional system. Then, we perform the measurement in the black box and basing on the measurement label, we measure the additional system. Eventually, basing on the outcome of the last measurement, we decide which of these measurements was contained in the black box.

All three approaches to the discrimination problem are first studied in the single-shot scenario, when the black box containing one of two measurements can be accessed only once. We also study the parallel discrimination scheme as well as the most general – adaptive scheme. The adaptive discrimination scheme allows for performing processing between subsequent queries to the black box, thus modifying the input for the subsequent query. The thesis of this dissertation concerns both multiple-shot discrimination schemes and states: *Adaptive schemes can improve the discrimination of quantum measurements.*

This dissertation consists of seven chapters and two appendices. Chapter 1 provides a general introduction and motivation. Mathematical preliminaries, basic

notions of quantum information theory and various useful mathematical tools are introduced in Chapter 2. Symmetric discrimination of quantum measurements is studied in single-shot and multiple-shot cases in Chapters 3 and 4, respectively. Unambiguous discrimination of quantum measurements is explored in Chapter 5. The following Chapter 6 is devoted to asymmetric discrimination. Final remarks and conclusions can be found in Chapter 7. There are also two appendices, which provide proofs which were too long and technical to be contained in the main text.

Chapter 1

Introduction

The problem of discrimination of various objects lies at the center of interest in many areas of technical and theoretical informatics. Machine learning algorithms are spreading far and wide and are currently inextricable elements of everyday life. Their applications range from banking, medicine and science and go far beyond that. Although machine learning algorithms are rather goal-oriented than abstract theoretical notions, they strongly rely on mathematics, mostly statistics, linear algebra and logic. One of the most common distinctions between machine learning algorithms differentiates two types of algorithms – supervised and unsupervised [1].

As far as unsupervised learning is concerned, its primary goal is to find the relationships and similarities in the dataset. Elements of the given dataset are described by some attributes. Basing on these attributes, the algorithm is supposed to aggregate the elements of the dataset into clusters with similar properties. In other words, the main task of unsupervised learning is splitting the dataset into classes of similar objects however, these algorithms can also be useful to find anomalies in the datasets. This task is called clustering, and it is often based on such algorithms as k-mean [2–4], k-nearest neighbors [5, 6], principal component analysis [7, 8].

Let us now elaborate a bit on the supervised learning. We are given a dataset with a few types of objects. Each object has some features and a label. The features describe the object and are usually written as vectors of some quantities. If the considered object was, e.g. a car, one could consider its features such as price, year of production, engine type, number of seats etc. Basing on these features, an experienced dealer can decide whether a car is worth buying or not. Therefore one can assign a car with a label describing whether the given car is a good value for money. Another classic example considers classifying customers who want to take a loan. Every customer has such features as their age, level of income, loan history etc. Basing on these features, the banker can decide whether a customer can get the loan or not. Thus the customer is given a label indicating if they obtained a

desired loan.

In supervised learning, we consider a set of objects, for instance, cars or bank customers. Each object is described by a vector of attributes (features) and a label stating to which class it belongs. This set is divided into the training and validation sets. The algorithm first needs to learn the properties of the classes, that is, to find the similarities among objects having the same labels. This is done on the training set, and this process is known as training the algorithm. Then, the validation set is taken into consideration. We give the algorithm some new input from the validation set, this time without the label, and ask the algorithm to predict the label basing on the knowledge gained during the training. To check the correctness of the algorithm, we can compare the resulting predicted labels with the known true labels. Basing on the number of correctly chosen labels, one is able to assess the quality of the classification rule. The typical examples of supervised learning algorithms include linear regression [9,10], logistic regression [11,12], Bayes classification [13,14], decision trees [15,16] and neural networks [17,18].

Now we provide a mathematical description of the basic classification problem. Assume our task is the classification of m classes of objects. Let $L = \{l_1, \ldots, l_m\}$ be a set of labels, where each label corresponds to a class. Let the dataset of objects to be classified be denoted by S. The elements of this set are represented as d-dimensional feature vectors. Therefore, i-th element of the dataset S takes the form $x_i = [x_{i,1}, \ldots, x_{i,d}]$. Every such vector has a corresponding label which indicates the class which this vector belongs. The goal of classification is to find a map $f : S \to L$, which associates each vector of attributes to a label. A typical solution for the classification problem concerns finding the hyperplane which splits the space into parts, where, in each of these parts, the elements have similar properties. In the most basic, two-dimensional case, the goal is to find a separating hyperplane which can distinguish between classes of different elements as well as possible. In a more general case, we allow for having more than two classes and improving the separating hyperplane using kernel methods.

In the typical formulation of the classification problem, when the model is already trained, we give it a new element and ask for its label. This classification problem can be seen also from a bit different perspective, and here we focus on another interpretation of this task. We will be interested in the problem of classification of m classes. Assume there are m devices, where i-th device produces elements from i-th class. Therefore, we can think about these elements as the outcomes of some device that produces elements having similar properties. Hence, the problem of classification between elements belonging to different classes can be translated to the problem of discrimination between various devices producing these elements. Now we arrive at the objective of this dissertation – the discrimination scheme. One of these devices is secretly chosen and put into a black box. We want

to find out which device is hidden in the black box, but we cannot just open it. We can see only the elements this device produces, not the device itself. There are many possible approaches towards this problem, including machine learning algorithms. We can use a supervised learning algorithm and train it to get some knowledge about the properties of the elements produced by each device. In the final discrimination scheme, we can simply use the trained algorithm to classify the given object and in this way decide which device was hidden in the black box. Nevertheless, basing only on the outputs of a certain device, we will usually not be able to state which device was hidden in the black box with certainty. It is crucial to have a reasonable measure of distance between the devices. If the devices produce elements according to some probability distributions, this, in fact, reduces to studying the distance between those probability distributions. An example of such a measure can be the well-established total variational distance between probability distributions.

The situation gets even more interesting when the device hidden in the black box not only produces elements, but rather transforms the given input into some output object. The description of how the devices act is known. The only thing we do not know is which of the devices is contained in the black box. In this case, to discriminate the devices, it is reasonable to think about what inputs should be used to get easily distinguishable outputs. In other words, we can use the knowledge about the devices to optimize the input, thus increasing the chances of correct discrimination.

Now we proceed to considering a quantum version of the discrimination problem. There are a few versions of the discrimination problem, which are discrimination of quantum states, channels and measurements. The simplest case that is the discrimination of quantum states, can be seen as discrimination of devices which only produce the quantum states. More precisely, we are given one of the devices, and each device produces only one quantum state. Therefore, we need to measure the resulting output state by a quantum measurement, and in this way, we obtain some classical information about the state. The problem of discrimination of quantum states is very well-researched [19–25]. An important property of discrimination of quantum states is that we are not always able to distinguish them perfectly. In fact, one can perfectly discriminate quantum states if and only if they are orthogonal.

Quantum channels can be seen as devices which transform one quantum state into another. Therefore, when discriminating quantum channels, one should take into consideration also the optimization of the input state. As the discrimination of quantum channels will be of significant interest in this dissertation, let us take a closer look at this scheme. There are two quantum channels, and we know their classical descriptions. One of them is secretly chosen and put into the black box. We can prepare any input state, so we look for a state that will result in as different

outputs as possible depending on which channel was hidden in the black box. Then, we apply the channel contained in the black box to the prepared input state, and, as a result, we obtain an output state. This resulting output is also a quantum state; hence we need to measure this state to get some classical information about it. Basing on the result of the final measurement, we make a decision which of the two channels was hidden in the black box.

As we are considering the discrimination of quantum objects, we can try to take advantage of quantum entanglement. More precisely, we can prepare an entangled input state on a larger space and apply the channel in the black box only on one part of the state. Finally, we measure the resulting state by a quantum measurement and basing on its classical output, we decide which of the channel was hidden in the black box. In many cases, this procedure substantially improves the discrimination [26–28].

Quantum measurements allow us to get classical information about quantum states and are typically used at the end of the procedure of discrimination both states and channels. Nevertheless, we can also consider the situation when the black box contains one of a few quantum measurements, which classical descriptions we do know. This problem, when the black box contains a secretly chosen quantum measurement, is the focus of attention in this dissertation. In the most basic discrimination scheme, we prepare some input state and measure it by the measurement contained in the black box. As a result, we obtain a classical label and we should make a decision which of the measurements was in the black box. This basic approach reduces to discrimination of probability distributions and it does not take any advantage of quantum entanglement. A more sophisticated version of the discrimination scheme allows for preparing an entangled input state on a compound register. Then, one part of this state is measured by the measurement contained in the black box. Finally, we measure the other part of the state by any prepared quantum measurement. Basing on the outcome of this final measurement, we make a decision which of the measurements was secretly chosen.

How to assess how good the discrimination was? We may be interested in verifying whether a specific object was given, and we want to be sure of that. We may also be satisfied with the situation when we know it only up to some probability. It may also happen that we want to avoid making a mistake in discrimination so much that we agree on the possibility of obtaining an inconclusive answer. There are three basic approaches towards discrimination that will be studied in this dissertation, which are symmetric discrimination, unambiguous discrimination and asymmetric discrimination. In the first symmetric approach, also known as minimum error discrimination, our goal is to decide which of the given devices was hidden in the black box, and we want to know it with as good probability as possible. It also means that whenever we make a decision which of the two devices

was chosen, we know it only up to some probability. The second approach assumes that when we get a conclusive result (that is, when we decide which of the devices was in the black box), we know it with certainty. It may happen, however, that we are not able to make a conclusive decision. In other words, if we know which device was hidden in the black box, we know it with probability one, but there is a chance that we will get an inconclusive answer. The third approach is also known as certification and it utilizes the statistical hypotheses testing. More precisely, we assume that one of the devices was in the black box and try to verify this hypothesis. In this approach, we differentiate two types of error, which are false positive and false negative errors and study them separately. We will be interested for example, when one type of error can be equal to zero and how small can be one type of error assuming some bound on the other type.

After using the black box in the discrimination procedure only once, we may not always be able to obtain satisfactory results of the discrimination. A natural solution to this problem includes using the quantum channel or measurement contained in the black box many times in various configurations. The most natural extension of the single-shot scheme is the parallel discrimination scheme. One can also try to make use of extra processing between queries to the black box. The latter approach is known as the adaptive discrimination scheme. In this dissertation, we will devote much attention to studying parallel and adaptive discrimination strategies. Essentially, the thesis of this dissertation can be written as:

Adaptive strategies can improve discrimination of quantum measurements.

In the literature, it is also common to study the case when the number of queries to the black box tends to infinity. The most popular approaches towards asymptotic discrimination of quantum channels include Chernoff [29], Heoffding [30] and Han-Kobayashi [31] settings. A broad introduction of all of these settings can be found, eg. in the work [32]. However, in this dissertation, we will restrict our attention to the situation when the black box can be used a finite number of times.

This dissertation is organized as follows. We begin with introducing basic notions of quantum information theory and mathematical preliminaries in Chapter 2.

Symmetric discrimination in the single-shot case is studied in Chapter 3. It is based on the works [33, 34] and focuses on the symmetric discrimination of two classes of quantum measurements: projective von Neumann measurements and symmetric informationally complete measurements, which are known as SIC POVMs. Multiple-shot discrimination is studied in Chapter 4. In this chapter, we will be interested in both parallel and adaptive discrimination strategies and we will explore the problem of when the adaptive strategy can outperform the parallel one. This chapter will be based on the work [35].

The following Chapter 5 concerns unambiguous discrimination. It is based on the work [35]. We will calculate the probability of unambiguous discrimination of

general measurements with rank-one effects. We will see how this probability can be written for the cases of discrimination of von Neumann measurements and SIC POVMs. We will generalize these results for the parallel case, and compare it with the performance of the adaptive scheme.

Asymmetric discrimination scheme is studied in Chapter 6. It contains the results from the works [36, 37]. We will prove the conditions when it is possible to exclude false negative error after a finite number of queries to the black box. We will also see if the adaptive scheme improves asymmetric discrimination. Finally, we will focus on the discrimination of von Neumann measurements and SIC POVMs in both single-shot and parallel schemes.

Conclusions and final remarks can be found in Chapter 7. There are also two appendices which contain proof of two theorems which require additional lemmas, and thus they were too long and technical to be stated in the main text.

Chapter 2

Preliminaries

2.1 Basic notions of quantum information

Let $\mathcal{X}$ be a complex Euclidean space. We will restrict our attention to finite-dimensional spaces so we can assume that $\mathcal{X} = \mathbb{C}^d$. The dimension of the space $\mathcal{X}$ will be denoted $\dim(\mathcal{X})$. Throughout this dissertation we will use the Dirac notation and write $|\phi\rangle \in \mathcal{X}$ for a column vector. Such a vector will be called a *ket*. The dual vector corresponding to the ket $|\phi\rangle$ will ce called a *bra* and denoted $\langle\phi| = |\phi\rangle^\dagger \coloneqq \overline{|\phi\rangle}^\top$. The symbol $(\cdot)^\dagger$ denotes the Hermitian conjugate of a vector, that is the transposition of the vector and complex conjugate of its elements.

For numerous applications it will be very convenient to distinguish a *canonical basis* of the space $\mathcal{X} = \mathbb{C}^d$. This basis will be denoted $\{|1\rangle, \ldots, |d\rangle\}$, there $|i\rangle$ is a vector which i-th entry is equal to one and all other entries are equal zero.

Let ϕ_i denote the i-th entry of $|\phi\rangle$. The inner product of vectors $|\phi\rangle, |\psi\rangle \in \mathcal{X}$ is defined as

$$\langle\phi|\psi\rangle \coloneqq \sum_{i=1}^{d} \overline{\phi_i}\psi_i. \tag{2.1}$$

This inner product $\langle\phi|\psi\rangle$ directly satisfies the following properties

1. linearity in the second argument, that is $\langle\phi|\alpha\psi_1 + \beta\psi_2\rangle = \alpha\langle\phi|\psi_1\rangle + \beta\langle\phi|\psi_2\rangle$ for every $\alpha, \beta \in \mathbb{C}$ and $|\phi\rangle, |\psi_1\rangle, |\psi_2\rangle \in \mathcal{X}$;
2. conjugate symmetry, that is $\langle\phi|\psi\rangle = \overline{\langle\psi|\phi\rangle}$ for all $|\phi\rangle, |\psi\rangle \in \mathcal{X}$;
3. positive definiteness, that is $\langle\phi|\phi\rangle \geq 0$ for all $|\phi\rangle \in \mathcal{X}$ and $\langle\phi|\phi\rangle = 0$ if and only if $|\phi\rangle = 0$.

The norm of the vector $|\phi\rangle$ is defined as $\||\phi\rangle\| \coloneqq \sqrt{\langle\phi|\phi\rangle} \geq 0$. One can also define the *outer product* of $|\psi\rangle$ and $|\phi\rangle$ as $|\psi\rangle\langle\phi|$ which corresponds to combining ket and bra together, to obtain a linear operator.

Operator-matrix correspondence A set of linear operators $A : \mathcal{X} \to \mathcal{Y}$, where $\mathcal{Y}$ is another complex Euclidean space, will be denoted by $\mathcal{L}(\mathcal{X}, \mathcal{Y})$. For the sake of simplicity we will write $\mathcal{L}(\mathcal{X})$ instead of $\mathcal{L}(\mathcal{X}, \mathcal{X})$.

Every linear operator $A \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$ can be associated with a matrix of size $\dim(\mathcal{Y}) \times \dim(\mathcal{X})$ with coefficients defined as $A_{i,j} \coloneqq \langle i|A|j\rangle$, where $|i\rangle \in \mathcal{Y}, |j\rangle \in \mathcal{X}$. Thanks to this operator-matrix correspondence, in the rest of this dissertation the notions of operators and matrices will be used interchangeably.

The symbol $\mathbb{1}_{\mathcal{X}}$ will be used to denote the identity matrix in the space $\mathcal{L}(\mathcal{X})$, that is a matrix which has ones on the diagonal and zeros everywhere else. We will neglect the index when it will be clear from the context.

Entry-wise conjugate, transpose and adjoint We will use the notation $\overline{A}$ to denote the entry-wise conjugate, that is a matrix such that $\overline{A}_{i,j} = \overline{A_{i,j}}$. The transposition of a matrix A will be denoted by $A^\top$. In other words, $A^\top_{i,j} = A_{j,i}$. Finally, we will use the notation $A^\dagger \coloneqq \overline{A^\top}$ for the adjoint of the matrix A.

Linear maps Let $\mathcal{X}$ and $\mathcal{Y}$ be complex Euclidean spaces. The set of linear maps $\Phi : \mathcal{L}(\mathcal{X}) \to \mathcal{L}(\mathcal{Y})$ will be denoted $\mathcal{T}(\mathcal{X}, \mathcal{Y})$. For the identity map we will use the notation $\mathbb{1}_{\mathcal{L}(\mathcal{X})} : \mathcal{L}(\mathcal{X}) \to \mathcal{L}(\mathcal{X})$ and we will neglect the index when the space will be clear from the context.

Tensor product Let $\mathcal{X} = \mathbb{C}^{d_1}$ and $\mathcal{Y} = \mathbb{C}^{d_2}$. The tensor products of vectors $|\phi\rangle \in \mathcal{X}$ and $|\psi\rangle \in \mathcal{Y}$ is defined as

$$|\phi\rangle \otimes |\psi\rangle \coloneqq \begin{bmatrix} \phi_1|\psi\rangle \\ \vdots \\ \phi_{d_1}|\psi\rangle \end{bmatrix}. \tag{2.2}$$

We will often write $|\phi\rangle|\psi\rangle$, or even $|\phi\psi\rangle$, instead of $|\phi\rangle \otimes |\psi\rangle$ to keep the notation short and concise.

The tensor product of spaces $\mathcal{X}$ and $\mathcal{Y}$ is denoted by $\mathcal{X} \otimes \mathcal{Y}$, and defined as

$$\mathcal{X} \otimes \mathcal{Y} \coloneqq \operatorname{span}\{|\phi\rangle \otimes |\psi\rangle : \ |\phi\rangle \in \mathcal{X}, |\psi\rangle \in \mathcal{Y}\}. \tag{2.3}$$

Such a product space will be also called a *compound* space. The spaces $\mathcal{X}$ and $\mathcal{Y}$ will be called the corresponding *registers* of this space. When talking about tensor products of many spaces we will use the notation $\mathcal{X}^{\otimes N}$ instead of writing N-time tensor product $\mathcal{X} \otimes \ldots \otimes \mathcal{X}$.

So far we introduced the tensor product of vectors and complex Euclidean spaces. Now we proceed to introducing tensor product of operators and linear

maps. Let $X \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$ and $Y \in \mathcal{Z}(\mathcal{Z}, \mathcal{W})$. The $X \otimes Y \in \mathcal{L}(\mathcal{X} \otimes \mathcal{Z}, \mathcal{Y} \otimes \mathcal{W})$ is defined as a unique operator satisfying

$$(X \otimes Y)\,|\phi\rangle|\psi\rangle = X|\phi\rangle \otimes Y|\psi\rangle \tag{2.4}$$

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

Let $\Phi \in \mathcal{T}(\mathcal{X}, \mathcal{Y})$ and $\Psi \in \mathcal{T}(\mathcal{Z}, \mathcal{W})$. The $\Phi \otimes \Psi \in \mathcal{T}(\mathcal{X} \otimes \mathcal{Z}, \mathcal{Y} \otimes \mathcal{W})$ is defined as a unique map satisfying

$$(\Phi \otimes \Psi)\,(X \otimes Y) = \Phi(X) \otimes \Psi(Y) \tag{2.5}$$

for every $X \in \mathcal{L}(\mathcal{X}),\ Y \in \mathcal{L}(\mathcal{Z})$.

Rank of an operator Let $A \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$. The subspace of $\mathcal{Y}$, called the *image* of A, is defined as

$$\mathrm{im}(A) \coloneqq \{A|x\rangle :\ |x\rangle \in \mathcal{X}\}. \tag{2.6}$$

The *rank* of the operator A is equal to the dimension of $\mathrm{im}(A)$, that is

$$\mathrm{rank}(A) = \dim(\mathrm{im}(A)). \tag{2.7}$$

Eigenvalues and eigenvectors Let $A \in \mathcal{L}(\mathcal{X})$. A complex number λ, such that $A|u\rangle = \lambda|u\rangle$ for some vector $|u\rangle \in \mathcal{X}$, is called an *eigenvalue* of the matrix A and the vector $|u\rangle$ is called an *eigenvector* of the matrix A. The multiset of eigenvalues of A is known as its *spectrum* and denoted $\mathrm{spec}(A)$.

Singular values and vectors While eigenvalues and eigenvectors are defined only for square matrices, in many cases it is useful to take advantage of their generalizations into singular values and vectors, which are defined for arbitrary matrices. Let $A \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$ have rank equal k. Then, from the *Singular Value Theorem* [38], one can write the matrix A as

$$A = \sum_{i=1}^{k} s_i |y_i\rangle\langle x_i|, \tag{2.8}$$

where $\{|x_1\rangle, \ldots, |x_k\rangle\} \subset \mathcal{X}$ and $\{|y_1\rangle, \ldots, |y_k\rangle\} \subset \mathcal{Y}$ are sets of orthonormal vectors and $s_1, \ldots, s_k$ are positive numbers.

The numbers $s_1, \ldots, s_k$ are called *singular values of the matrix* A. The vectors $\{|x_1\rangle, \ldots, |x_k\rangle\}$ are called *right singular vectors* and vectors $\{|y_1\rangle, \ldots, |y_k\rangle\}$ are called *left singular vectors* of the matrix A.

Trace and partial trace A *trace* of a matrix $A \in \mathcal{L}(\mathcal{X})$ is defined as a unique mapping $\mathrm{Tr} : \mathcal{L}(\mathcal{X}) \to \mathbb{C}$ satisfying

$$\mathrm{Tr}(|\phi\rangle\langle\psi|) = \langle\psi|\phi\rangle \tag{2.9}$$

for every $|\phi\rangle, |\psi\rangle \in \mathcal{X}$. It can be calculated as a sum if its diagonal elements, that is $\mathrm{Tr}(A) := \sum_i A_{i,i}$.

The *cyclic property* of the trace, which will be crucial in numerous proofs, can be formulated as $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$ for every $A \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$ and $B \in \mathcal{L}(\mathcal{Y}, \mathcal{X})$. Another important property of the trace states that it is in fact a sum of the eigenvalues of the matrix.

The linear map $\mathrm{Tr}_{\mathcal{X}} : \mathcal{L}(\mathcal{X} \otimes \mathcal{Y}) \to \mathcal{L}(\mathcal{Y})$ is called a *partial trace* and is defined as a unique map satisfying

$$\mathrm{Tr}_{\mathcal{X}}\left(X \otimes Y\right) = \left(\mathrm{Tr} \otimes \mathbb{1}_{\mathcal{L}(\mathcal{Y})}\right)\left(X \otimes Y\right) = \mathrm{Tr}(X)Y \tag{2.10}$$

for every $X \in \mathcal{L}(\mathcal{X})$ and $Y \in \mathcal{L}(\mathcal{Y})$. Similarly, a linear map $\mathrm{Tr}_{\mathcal{Y}} : \mathcal{L}(\mathcal{X}\otimes\mathcal{Y}) \to \mathcal{L}(\mathcal{X})$ is defined as a unique map satisfying

$$\mathrm{Tr}_{\mathcal{Y}}\left(X \otimes Y\right) = \left(\mathbb{1}_{\mathcal{L}(\mathcal{X})} \otimes \mathrm{Tr}\right)\left(X \otimes Y\right) = \mathrm{Tr}(Y)X \tag{2.11}$$

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

Vectorization Given complex Euclidean spaces $\mathcal{X}$ and $\mathcal{Y}$, the (lexicographic) *vectorization* is a linear map $\mathrm{vec} : \mathcal{L}(\mathcal{Y}, \mathcal{X}) \to \mathcal{X} \otimes \mathcal{Y}$ defined by the action on the elements of the canonical bases

$$\mathrm{vec}\left(|i\rangle\langle j|\right) = |i\rangle \otimes |j\rangle, \tag{2.12}$$

where $|i\rangle \in \mathcal{X}, |j\rangle \in \mathcal{Y}$.

From the linearity, the action of vectorization can be generalized for any for $|u\rangle \in \mathcal{X}, |v\rangle \in \mathcal{Y}$ as

$$\mathrm{vec}\left(|u\rangle\langle v|\right) = |u\rangle \otimes \overline{|v\rangle}. \tag{2.13}$$

In the rest of this dissertation the vectorization of the matrix A will be denoted by $|A\rangle\!\rangle := \mathrm{vec}(A)$. Moreover, $\langle\!\langle A| := (|A\rangle\!\rangle)^\dagger$.

Two properties of vectorization, which hold for all $A, B \in \mathcal{L}(\mathcal{Y}, \mathcal{X})$ will be of great advantage in various proofs, which are [38]

$$\mathrm{Tr}_{\mathcal{X}}\left(|A\rangle\!\rangle\langle\!\langle B|\right) = A^\top\overline{B}, \quad \mathrm{Tr}_{\mathcal{Y}}\left(|A\rangle\!\rangle\langle\!\langle B|\right) = AB^\dagger. \tag{2.14}$$

Let now $A \in \mathcal{L}(\mathcal{X}, \mathcal{Y}), B \in \mathcal{L}(\mathcal{Z}, \mathcal{W})$ and $X \in \mathcal{L}(\mathcal{Z}, \mathcal{X})$. The important

property of the vectorization, known as *telegraphic notation* [38], is given by

$$(A \otimes B)\,|X\rangle\!\rangle = |AXB^{\top}\rangle\!\rangle. \tag{2.15}$$

We will often also use the operation which is a reverse of vectorization. This operation is a linear map $\mathrm{vec}^{\dagger} : \mathcal{X} \otimes \mathcal{Y} \to \mathcal{L}(\mathcal{Y}, \mathcal{X})$ defined by the action on the elements of the canonical bases

$$\mathrm{vec}^{\dagger}\left(|i\rangle \otimes |j\rangle\right) = |i\rangle\langle j|, \tag{2.16}$$

where $|i\rangle \in \mathcal{X}, |j\rangle \in \mathcal{Y}$. We will often use the simplified notation $[\psi]$ to denote $\mathrm{vec}^{\dagger}(|\psi\rangle)$, where $|\psi\rangle \in \mathcal{X} \otimes \mathcal{Y}$.

Classes of operators which will be studied in this dissertation

- An operator $A \in \mathcal{L}(\mathcal{X})$ is called *normal* if $AA^{\dagger} = A^{\dagger}A$.
- An operator $A \in \mathcal{L}(\mathcal{X})$ is called *Hermitian* if $A = A^{\dagger}$. The set of Hermitian operators will be denoted by $\mathrm{Herm}(\mathcal{X})$. The eigenvalues of Hermitian matrix are real numbers.
- An operator $A \in \mathcal{L}(\mathcal{X})$ is called *positive semidefinite*, and denoted $A \geq 0$, when $\langle\psi|A|\psi\rangle \geq 0$ for every $|\psi\rangle \in \mathcal{X}$. Alternatively, the set of positive semidefinite operators, denoted $\mathrm{Pos}(\mathcal{X})$, can be defined as the subset of all Hermitian operators which all eigenvalues are non-negative.
- Similarly, $A \in \mathcal{L}(\mathcal{X})$ is called *positive definite* when $\langle\psi|A|\psi\rangle > 0$ for every $|\psi\rangle \in \mathcal{X}$. The set of positive definite operators will be denoted $\mathrm{Pd}(\mathcal{X})$ and we will write $A > 0$ when a matrix A is positive definite. Alternatively, we can say that a matrix is positive definite if it is Hermitian and all its eigenvalues are positive.
- An operator $A \in \mathrm{Pos}(\mathcal{X})$ is called a *projection operator* if $A^2 = A$. The set of projection operators will be denoted $\mathrm{Proj}(\mathcal{X})$. Alternatively, a projection operator can be defined as a Hermitian operator having only eigenvalues equal either zero or one.
- A matrix $X \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$ is called an *isometry* if $\dim(\mathcal{X}) \leq \dim(\mathcal{Y})$ and $XX^{\dagger} = \mathbb{1}_{\mathcal{X}}$. The set of isometry operators will be denoted $U(\mathcal{X}, \mathcal{Y})$. An isometry is a rectangular matrix when $\dim(\mathcal{X}) \neq \dim(\mathcal{Y})$. A square isometry is called a *unitary operator* and the set of unitary operations will be denoted $U(\mathcal{X})$. An alternative definition for unitary operators says that $U \in U(\mathcal{X})$ if $UU^{\dagger} = U^{\dagger}U = \mathbb{1}_{\mathcal{X}}$.

The notation of matrix inequalities $A \geq B$ ($A > B$) should be understood as $A - B \geq 0$ ($A - B > 0$).

Spectral decomposition and functions of normal matrices Let $A \in \mathcal{L}(\mathcal{X})$ be a normal operator of dimension d having eigenvalues $\lambda_1, \ldots, \lambda_d$. From the *Spectral Theorem* [38] we can write A as

$$A = \sum_{i=1}^{d} \lambda_i |x_i\rangle\langle x_i|, \tag{2.17}$$

where $\{|x_1\rangle, \ldots, |x_d\rangle\}$ is an orthonormal basis of $\mathcal{X}$ and elements of this basis are called *eigenvectors*.

In this dissertation we will be only interested in functions acting on the sets of normal operators. Using the notation as above, a function $f : \mathcal{L}(\mathcal{X}) \to \mathcal{L}(\mathcal{X})$ is defined as

$$f(A) = \sum_{i=1}^{d} f(\lambda_i) |x_i\rangle\langle x_i|. \tag{2.18}$$

Jordan-Hahn decomposition Let $H \in \mathrm{Herm}(\mathcal{X})$. From Spectral Theorem and the property of Hermitian operator we know that H can be written as $H = \sum_{i=1}^{d} \lambda_i |x_i\rangle\langle x_i|$, where all $\lambda_1, \ldots, \lambda_d$ are real numbers. To formulate the Jordan-Hahn decomposition of H, we define operators

$$\begin{aligned} P &:= \sum_{i=1}^{d} \max\{\lambda_i, 0\} |x_i\rangle\langle x_i|, \\ Q &:= \sum_{i=1}^{d} \max\{-\lambda_i, 0\} |x_i\rangle\langle x_i|, \end{aligned} \tag{2.19}$$

which are called the *non-negative* and *non-positive* parts, respectively. The expression

$$H = P - Q, \tag{2.20}$$

where $P, Q \in \mathrm{Pos}(\mathcal{X})$ and $PQ = 0$, is known as *Jordan-Hahn decomposition* of the Hermitian operator H [38]. The above requirements for P and Q assure that there is only one possible such decomposition, hence the Jordan-Hahn decomposition is unique.

Norms of operators Throughout this dissertation we will often use Schatten p-norm. Let $A \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$. For any number $p \geq 1$ one defines the Schatten p-norm

of A as

$$\|A\|_p \coloneqq \left(\mathrm{Tr}\left(\left(A^\dagger A\right)^{\frac{p}{2}}\right)\right)^{\frac{1}{p}} \tag{2.21}$$

and

$$\|A\|_\infty \coloneqq \lim_{p\to\infty} \|A\|_p = \max\left\{\|A|u\rangle\| : \ |u\rangle \in \mathcal{X}, \||u\rangle\| \leq 1\right\}. \tag{2.22}$$

The Schatten 1-norm is also known as the *trace norm*. It can be calculated directly as

$$\|A\|_1 \coloneqq \mathrm{Tr}\left(\sqrt{A^\dagger A}\right). \tag{2.23}$$

The trace norm coincides with the sum of singular values of the matrix X. The dual definition of the trace norm yields [38]

$$\|A\|_1 = \max_{V\in\mathcal{U}(\mathcal{X})} |\mathrm{tr}AV|. \tag{2.24}$$

The Schatten 2-norm is commonly called the *Frobenius norm*. It is analogous to Euclidean norm for vectors and it can be calculated as

$$\|A\|_2 \coloneqq \sqrt{\mathrm{Tr}\left(A^\dagger A\right)} = \||A\rangle\rangle\|. \tag{2.25}$$

The Schatten ∞-norm is the norm induced by the Euclidean norm and is equal to the largest singular value of the matrix. It is known in the literature as the *spectral norm* and we will often be writing $\|A\|$ instead of $\|A\|_\infty$ to denote the spectral norm of A.

Finally, let us quote a useful lemma from [38] which gives direct formulas for the eigenvalues and trace norm of rank-two matrices and can be obtained by direct calculations.

Lemma 1 *The operator $\alpha|u\rangle\langle u| - \beta|v\rangle\langle v|$, where $\||u\rangle\| = \||v\rangle\| = 1$ and $\alpha, \beta \in \mathbb{R}$, is Hermitian and has at most two nonzero eigenvalues given by the expression*

$$\lambda_\pm = \frac{\alpha-\beta}{2} \pm \frac{1}{2}\sqrt{(\alpha+\beta)^2 - 4\alpha\beta|\langle u|v\rangle|^2}. \tag{2.26}$$

Moreover

$$\|\alpha|u\rangle\langle u| - \beta|v\rangle\langle v|\|_1 = \sqrt{(\alpha+\beta)^2 - 4\alpha\beta|\langle u|v\rangle|^2}. \tag{2.27}$$

2.1.1 Quantum states

Finally, we are in position to introduce the definition of the set of quantum states. A quantum state represents a generalized probability distribution. In this subsection we will begin with the formal definition of quantum states and later, we will state some of their properties and decompositions.

Definition 1 *A quantum state $\rho \in \mathcal{L}(\mathcal{X})$ is a positive-semidefinite operator satisfying* $\mathrm{Tr}(\rho) = 1$. *The set of quantum states will be denoted by* $\mathcal{D}(\mathcal{X})$.

The set $\mathcal{D}(\mathcal{X})$ is also known an the set of *density operators* and it coincides with the definition of quantum states.

A significant class of quantum states are pure states. Formally, a quantum state $\rho \in \mathcal{D}(\mathcal{X})$ is called a *pure state* when $\mathrm{rank}(\rho) = 1$. From the Spectral Theorem [38] it follows that that the pure states are the extreme points of the convex set $\mathcal{D}(\mathcal{X})$. Every pure state can be written as a projector on 1-dimensional subspace, that is $|\psi\rangle\langle\psi|$, where $\||\psi\rangle\| = 1$, and every projector expressed in this way is a pure quantum state. While talking about pure quantum states we will often be writing only $|\psi\rangle$ instead of $|\psi\rangle\langle\psi|$. A quantum state which is not pure, will be called a *mixed* state.

Entanglement A quantum state $\rho \in \mathcal{D}(\mathcal{X} \otimes \mathcal{Y})$ is called *separable*, if there exist collections of states $\{\sigma_1, \ldots, \sigma_m\} \subseteq \mathcal{D}(\mathcal{X})$, $\{\phi_1, \ldots, \phi_m\} \subseteq \mathcal{D}(\mathcal{Y})$ and a probability vector $(p_1, \ldots, p_m)$ such that

$$\rho = \sum_{i=1}^{m} p_i \sigma_i \otimes \phi_i. \tag{2.28}$$

A quantum state, which is not separable, is called an *entangled* state.

Entanglement is a key resource in quantum information theory, which will be of great importance in the discrimination problems studied in this dissertation [26, 39]. The criteria for verifying if a quantum states are entangled were first proved in [40], and numerous results regarding for instance multipartite entanglement [41–43], geometry of entangled states [44, 45] and entanglement measures [46–50] were reported. The broad review of the results concerning quantum entanglement can be fund in [51].

Now we introduce two decompositions of quantum states, which are the Spectral Decomposition of mixed states and the Schmidt Decomposition of pure states on a compound space.

Spectral Decomposition Let $\rho \in \mathcal{D}(\mathcal{X})$ and $d = \dim(\mathcal{X})$. The quantum state ρ can be decomposed into a sum as in Eq. (2.17), where $\{|x_1\rangle, \ldots, |x_d\rangle\}$ is an orthonormal basis of $\mathcal{X}$ and $(\lambda_1, \ldots, \lambda_d)$ is a probability vector. The scalars λ_i are eigenvalues of the state ρ and vectors $|x_i\rangle$ are the corresponding eigenvectors.

Schmidt Decomposition Every pure quantum state $|\psi\rangle \in \mathcal{X} \otimes \mathcal{Y}$, on a compound space $\mathcal{X} \otimes \mathcal{Y}$, can be expressed as

$$|\psi\rangle = \sum_{i=1}^{r} \sqrt{\lambda_i} |x_i\rangle \otimes |y_i\rangle, \tag{2.29}$$

where $\{|x_i\rangle\}_i$ is an orthonormal basis of the space $\mathcal{X}$, $\{|y_i\rangle\}_i$ is an orthonormal basis of the space $\mathcal{Y}$ and $\lambda_1, \ldots, \lambda_r$ are positive real numbers. This decomposition is called the *Schmidt Decomposition* and it follows directly from the Singular Value Decomposition in Eq. (2.8) and the property of vectorization in Eq. (2.13). The number r is known as the *Schmidt rank*.

2.1.2 Quantum channels and linear maps

We say that a linear map $\Phi \in \mathcal{T}(\mathcal{X}, \mathcal{Y})$ is *positive* when for every $X \geq 0$ it holds that $\Phi(X) \geq 0$. We say that a linear map is *completely positive*, when for every $X \in \mathrm{Pos}(\mathcal{X} \otimes \mathcal{Z})$ and for every space $\mathcal{Z}$ it holds that $(\Phi \otimes \mathbb{1}_{\mathcal{L}(\mathcal{Z})})(X) \geq 0$. We say that a linear map $\Phi \in \mathcal{T}(\mathcal{X}, \mathcal{Y})$ is *trace-preserving*, when $\mathrm{Tr}(\Phi(X)) = \mathrm{Tr}(X)$ for every $X \in \mathcal{L}(\mathcal{X})$ and it is *Hermiticity-preserving* if for every $X \in \mathrm{Herm}(\mathcal{X})$ it holds that $\Phi(X) \in \mathrm{Herm}(\mathcal{Y})$.

Definition 2 *A* quantum channel *is a linear map $\Phi \in \mathcal{T}(\mathcal{X}, \mathcal{Y})$, which is completely positive and trace-preserving. The set of such quantum channels will be denoted by $\mathcal{C}(\mathcal{X}, \mathcal{Y})$.*

Important classes of quantum channels We will focus on two classes of quantum channel, which are isometry channels and dephasing channel.

- Let $U \in \mathcal{U}(\mathcal{X}, \mathcal{Y})$ be an isometry operator. The *isometry channel* $\Phi_U \in \mathcal{C}(\mathcal{X}, \mathcal{Y})$ is defined as
$$\Phi_U(\rho) = U \rho U^\dagger. \tag{2.30}$$
When $U \in \mathcal{U}(\mathcal{X})$ is a unitary matrix, then the channel Φ_U is called a *unitary channel*.

- The *completely dephasing channel* $\Delta \in \mathcal{C}(\mathcal{X})$, where $\dim(\mathcal{X}) = d$, is defined as
$$\Delta(\rho) = \sum_{i=1}^{d} \langle i|\rho|i\rangle |i\rangle\langle i|. \tag{2.31}$$

The set of linear maps $\mathcal{T}(\mathcal{X}, \mathcal{Y})$ has several representations. We will focus on the Kraus, Choi-Jamiołkowski and Stinespring representations.

Kraus representation Every linear map $\Phi \in \mathcal{T}(\mathcal{X}, \mathcal{Y})$ can be represented as

$$\Phi(\rho) = \sum_{i=1}^{k} E_i \rho F_i^\dagger, \tag{2.32}$$

where operators $\{E_i\}_{i=1}^k, \{F_i\}_{i=1}^k \subseteq \mathcal{L}(\mathcal{X}, \mathcal{Y})$ are called *Kraus operators* of the given map Φ [52–54].

In the case when $\Phi \in \mathcal{C}(\mathcal{X}, \mathcal{Y})$ is a quantum channel, then it holds that $E_i = F_i$ for every $i = 1, \ldots, k$ and $\sum_{i=1}^k E_i^\dagger E_i = \mathbb{1}_\mathcal{X}$. The former condition corresponds to complete positivity of the quantum channel while the latter assures that the channel is trace-preserving. Therefore, with notation as above, the Kraus representation of the quantum channels Φ can be written as

$$\Phi(\rho) = \sum_{i=1}^{k} E_i \rho E_i^\dagger. \tag{2.33}$$

Let us see the Kraus representation of exemplary quantum channels. Unitary channel Φ_U has the Kraus representation which consists of only one operator, that is $\{E_i\}_i = \{U\}$. The completely dephasing channel Δ has Kraus representation $\{|i\rangle\langle i|\}_{i=1}^d$

An important feature of Kraus operators yields that they are not unique. More precisely, consider two collections of operators $\{E_i\}_{i=1}^k$ and $\{F_i\}_{i=1}^k$ which for every ρ satisfy

$$\sum_{i=1}^{k} E_i \rho E_i^\dagger = \sum_{i=1}^{k} F_i \rho F_i^\dagger. \tag{2.34}$$

Then there exists a unitary operator $U \in \mathcal{U}(\mathcal{X})$ such that

$$F_i = \sum_{j=1}^{k} \langle i|U|j\rangle A_j \tag{2.35}$$

holds for every $i = 1, \ldots, k$.

Choi-Jamiołkowski representation The Choi-Jamiołkowski representation [53, 55] of $\Phi \in \mathcal{T}(\mathcal{X}, \mathcal{Y})$ is defined as a mapping $J : \mathcal{T}(\mathcal{X}, \mathcal{Y}) \to \mathcal{L}(\mathcal{Y} \otimes \mathcal{X})$ as

$$J(\Phi) = \sum_{i,j} \Phi(|i\rangle\langle j|) \otimes |i\rangle\langle j| = \left(\Phi \otimes \mathbb{1}_{\mathcal{L}(\mathcal{X})}\right)(|\mathbb{1}_\mathcal{X}\rangle\rangle\langle\langle \mathbb{1}_\mathcal{X}|). \tag{2.36}$$

The rank of the Choi-Jamiołkowski operator $J(\Phi)$ is called the *Choi rank* of Φ. The action of the linear map Φ on a state ρ can be recovered as

$$\Phi(\rho) = \mathrm{Tr}_{\mathcal{X}} \left(J(\Phi) \left(\mathbb{1}_{\mathcal{Y}} \otimes \rho^{\top} \right) \right). \tag{2.37}$$

This representation is particularly important as is resembles significant algebraic properties. The map Φ is Hermiticity-preserving if and only if $J(\Phi) \in \mathrm{Herm}(\mathcal{Y} \otimes \mathcal{X})$. Moreover, Φ is completely positive if and only if $J(\Phi) \in \mathrm{Pos}(\mathcal{Y} \otimes \mathcal{X})$. Finally, Φ is trace-preserving if and only if $\mathrm{Tr}_{\mathcal{Y}} \left(J(\Phi) \right) = \mathbb{1}_{\mathcal{X}}$.

The Choi-Jamiołkowski representation of a unitary channel Φ_U has the form $J\left(\Phi_U\right) = |U\rangle\!\rangle\langle\!\langle U|$.

Stinespring representation Let $\mathcal{X}, \mathcal{Y}, \mathcal{Z}$ be complex Euclidean spaces. For $\Phi \in \mathcal{T}(\mathcal{X}, \mathcal{Y})$, its Stinespring representation is defined for every $\rho \in \mathcal{D}(\mathcal{X})$ as

$$\Phi(\rho) = \mathrm{Tr}_{\mathcal{Z}} \left(A \rho B^{\dagger} \right), \tag{2.38}$$

where $A, B \in \mathcal{L}(\mathcal{X}, \mathcal{Y} \otimes \mathcal{Z})$ [56]. Stinespring representation of a quantum map is not unique. In the case when $\Phi \in \mathcal{C}(\mathcal{X}, \mathcal{Y})$, its Stinespring representation can be written as

$$\Phi(\rho) = \mathrm{Tr}_{\mathcal{Z}} \left(V \rho V^{\dagger} \right) \tag{2.39}$$

for some isometry matrix $V \in \mathcal{U}(\mathcal{X}, \mathcal{Y} \otimes \mathcal{Z})$.

Moreover, the Stinespring representation of $\Phi \in \mathcal{C}(\mathcal{X})$ can be written by the use of unitary operator as

$$\Phi(\rho) = \mathrm{Tr}_{\mathcal{Z}} \left(U \left(\rho \otimes |0\rangle\langle 0| \right) U^{\dagger} \right), \tag{2.40}$$

where $U \in \mathcal{U}(\mathcal{X} \otimes \mathcal{Z})$. This representation has a useful operational interpretation. We consider the action of a unitary channel on a given system with an attached additional state $|0\rangle\langle 0|$. Then, we perform the partial trace on the additional register and obtain the action of the original channel.

2.1.3 Quantum measurements

The most general quantum measurements are positive operator valued measures (POVMs). In this dissertation we will consider only the POVMs having a finite number of effects. Formally, a collection of is positive semidefinite operators $\mathcal{P} = \{E_1, \ldots, E_m\} \subset \mathrm{Pos}(\mathcal{X})$ is called a POVM if $\sum_{i=1}^{m} E_i = \mathbb{1}_{\mathcal{X}}$. The operators E_i are called *effects*. When a quantum state ρ is measured by the measurement $\mathcal{P}$, then the label i is obtained with probability $\mathrm{Tr}\left(\rho E_i\right)$ and the state ρ ceases to exist (Born rule).

Every quantum measurement $\mathcal{P} = \{E_1, \ldots, E_m\}$ can be identified with a measure-and-prepare channel which gives a classical output. In other words, the output of this channel is a diagonal state where the probability distribution on the diagonal gives the probabilities of obtaining the measurement labels. The action of this channel on a quantum state ρ can be written as

$$\mathcal{P}(\rho) = \sum_{i=1}^{m} \mathrm{Tr}(E_i \rho) |i\rangle\langle i|. \tag{2.41}$$

The Choi-Jamiołkowski representation of the quantum measurement $\mathcal{P}$ has a block-diagonal structure

$$J(\mathcal{P}) = \sum_{i=1}^{m} |i\rangle\langle i| \otimes E_i^\top, \tag{2.42}$$

which can be seen from direct calculation

$$\begin{aligned} J(\mathcal{P}) &= \sum_{k,l=1}^{n} \mathcal{P}(|k\rangle\langle l|) \otimes |k\rangle\langle l| = \sum_{k,l=1}^{n} \sum_{i=1}^{n} \mathrm{Tr}\left(|k\rangle\langle l| E_i\right) |i\rangle\langle i| \otimes |k\rangle\langle l| \\ &= \sum_{i=1}^{n} \sum_{k,l=1}^{n} |i\rangle\langle i| \otimes \langle k| E_i^\top |l\rangle |k\rangle\langle l| = \sum_{i=1}^{n} |i\rangle\langle i| \otimes E_i^\top. \end{aligned} \tag{2.43}$$

When the effects of the measurement are projection operators, then such a measurement is called the *projective measurement*. Throughout this dissertation, we will focus mostly on two classes of quantum measurements, which are projective von Neumann measurements and SIC POVMs, which are described below.

Von Neumann measurement

Let $\{|u_1\rangle, \ldots, |u_d\rangle\}$ be an orthonormal basis of the space $\mathcal{X}$. A quantum measurement with effects $\{|u_1\rangle\langle u_1|, \ldots, |u_d\rangle\langle u_d|\}$ is called a *von Neumann measurement*. Noting that $|u_i\rangle = U|i\rangle$ is the i-th column of some unitary matrix $U \in \mathcal{U}(\mathcal{X})$, we can parameterize the von Neumann measurement by this unitary matrix. Therefore, we will simply write $\mathcal{P}_U$ to denote the von Neumann measurement with effects $\{|u_1\rangle\langle u_1|, \ldots, |u_d\rangle\langle u_d|\}$.

In this definition of von Neumann measurement $\mathcal{P}_U$, every unitary matrix taken from the set $\{UE : \ E \in \mathcal{DU}(\mathcal{X})\}$, where $\mathcal{DU}$ denotes the set of diagonal unitary matrices, specifies the same measurement. It can be easily seen that a projection $U|i\rangle\langle i|U^\dagger$ built from unitary some matrix U will be the same as the projection built from the unitary matrix UE for some $E \in \mathcal{DU}(\mathcal{X})$. Hence, we an say that matrices

UE form an equivalence class of the unitary matrix U. In other words $\mathcal{P}_U = \mathcal{P}_{UE}$ for every $E \in \mathcal{DU}(\mathcal{X})$.

We will often take advantage of the von Neumann measurement in the canonical basis, that is the measurement associated with the identity matrix $\mathbb{1}$. Therefore the effects of this measurements are matrices of the form $\{|1\rangle\langle 1|, \ldots, |d\rangle\langle d|\}$. Note that such a measurement corresponds to the completely dephasing channel Δ, which was introduced in Eq. (2.31).

SIC POVMs

Another important class of quantum measurements are symmetric informationally complete (SIC) POVMs [57–61]. A SIC POVM of dimension d has d^2 effects $\{|x_1\rangle\langle x_1|, \ldots, |x_{d^2}\rangle\langle x_{d^2}|\}$, where $|x_i\rangle\langle x_i| = \frac{1}{d}|\phi_i\rangle\langle\phi_i|$ and $\||\phi_i\rangle\| = 1$ for every $i = 1, \ldots, d^2$. Moreover, the symmetry condition states that

$$|\langle\phi_i|\phi_j\rangle|^2 = \frac{1}{d+1}, \tag{2.44}$$

for $i \neq j$.

It is not known whether SIC POVMs exist for every dimension and it is also an open question whether it is possible to construct an infinite family of them [62–65]. Moreover, analytical construction of SIC POVMs is very complex and for many dimensions only numerical results prove the existence of them.

In further sections, we will study discrimination between two SIC POVMs of the same dimension. The effects of the second measurement will be a permutation of effects of the first measurement. Therefore, let us formally introduce the notion of permutations and notation.

A *permutation* π of a set $S \coloneqq \{1, \ldots, m\}$ is a bijection function from S to itself. Loosely speaking, it corresponds to rearrangement of elements. We will use the notation $\pi = (a_1, a_2 \ldots, a_m)$ instead of writing

$$\pi(1) = a_1, \pi(2) = a_2, \ldots, \pi(m) = a_m. \tag{2.45}$$

A useful property of a permutation in the number of fixed points. A *fixed point* of the permutation π is the number $a \in S$ for which $\pi(a) = a$. The number of fixed points of the permutation is the cardinality of the set of its fixed points, and will be denoted by $k \coloneqq \#\{a : \ \pi(a) = a\}$.

2.2 Distance between quantum objects

While studying the problem of discrimination of quantum objects, we will often need to be able to measure the distance between quantum objects. In this section we will introduce notions of distance between probability distributions, quantum states, channels and measurements.

2.2.1 Distance between probability distributions

Let p and q be two probability distributions on d dimensional space $\mathcal{X}$. The *total variation distance* between p and q is defined as

$$\|p - q\|_1 \coloneqq \sum_i |p_i - q_i| = 2 \max_{\Delta \subseteq \{1,\ldots,d\}} \sum_{a \in \Delta} (p_a - q_a) \,. \tag{2.46}$$

This notion of distance has an operational interpretation which states that it is the greatest possible difference between the probabilities that distributions p and q can assign to the same event. It also relates to the an error when using the maximum likelihood method. Due to this interpretation, the total variation distance between probability distributions is commonly used in machine learning and statistics.

2.2.2 Distance between quantum states

Trace distance Given two quantum states $\rho, \sigma \in \mathcal{D}(\mathcal{X})$, the *trace distance* between ρ and σ is defined as

$$\mathrm{dist}(\rho, \sigma) \coloneqq \|\rho - \sigma\|_1. \tag{2.47}$$

Note that when quantum states ρ and σ are diagonal, then calculating the trace distance between them reduces to calculating the total variation distance between probability distributions on their diagonals.

The trace norm is typically used to describe the distance between quantum states due to its operational interpretation. It is a natural generalization of the distance between probability distributions to the distance between quantum states. It it also used in the Holevo-Helstrom Theorem [66], which gives an upper bound on the probability of successful discrimination between quantum states.

Theorem 1 (Holevo-Helstrom) *Let $\mathcal{X}$ be a complex Euclidean space and $\rho, \sigma \in \mathcal{D}(\mathcal{X})$. For every measurement $\mathcal{P} = \{E_0, E_1\}$ and for every $\lambda \in [0, 1]$ it holds that*

$$\lambda \mathrm{tr}\,(E_0 \rho) + (1 - \lambda)\mathrm{tr}\,(E_1 \sigma) \leq \frac{1}{2} + \frac{1}{2} \|\lambda \rho - (1 - \lambda)\sigma\|_1. \tag{2.48}$$

Moreover, there exists a projective measurement $\mathcal{P}$ for which the inequality is saturated.

Now we will state the form of the projective measurement which saturates the bound in the above theorem. Let $X \coloneqq \lambda\rho - (1-\lambda)\sigma$ and let $X = P - Q$ be its Jordan-Hahn decomposition for $P, Q \in \mathrm{Pos}(\mathcal{X})$. The projective measurement is defined as

$$E_0 \coloneqq \mathrm{Proj}(\mathrm{im}(P)), \quad E_1 \coloneqq \mathbb{1} - \mathrm{Proj}(\mathrm{im}(P)). \tag{2.49}$$

Let us elaborate a bit on the operational interpretation of the Holevo-Helstrom theorem. We are given one of the two quantum states, either ρ or σ. Our task is to decide which of the states we were given. We know that the state ρ was given with probability λ and the state σ was given with probability $1-\lambda$. To perform the discrimination, we prepare the measurement $\mathcal{P}$ and measure the given state. The measurement $\mathcal{P}$ has two effects: E_0 and E_1. If the outcome of this measurement corresponds to the effect E_0, then we decide that the given state was ρ. Analogously, when the outcome of this measurement corresponds to the effect E_1, then we decide that the given state was σ.

The expression on the left hand side of Eq. (2.48) is the probability of successful discrimination between the states. First, we multiply the probability λ of being given the state ρ by the probability of obtaining the measurement label corresponding to the effect E_0. This part corresponds to making a correct discrimination decision when we were given the state ρ. Then, we multiply the probability of being given the state σ times the probability of obtaining the measurement label corresponding to the effect E_1. Finally, we sum the probabilities of correctly discriminating both ρ and σ. The Holevo-Helstrom theorem gives the upper bound on this probability expressed using the trace distance between quantum states. As this probability can be saturated by the use of measurement from Eq. (2.49), this theorem also provides the optimal strategy for discrimination of quantum states.

2.2.3 Distance between quantum channels

The most natural notion of distance between quantum channels is the diamond norm distance. For a linear map $\Phi \in \mathcal{T}(\mathcal{X}, \mathcal{Y})$, its *diamond norm* is defined as

$$\|\Phi\|_\diamond = \max_{\|X\|_1 \leq 1} \left\| \left(\Phi \otimes \mathbb{1}_{\mathcal{L}(\mathcal{X})}\right)(X) \right\|_1 . \tag{2.50}$$

When dealing with the diamond norm of a Hermiticity-preserving map, it may be beneficial to make use of the alternative formula for the diamond norm [38, 67]

$$\|\Phi\|_\diamond = \max\{\|(\mathbb{1} \otimes \sqrt{\rho}) J(\Phi) (\mathbb{1} \otimes \sqrt{\rho})\|_1 : \rho \in \mathcal{D}(\mathcal{X})\}. \tag{2.51}$$

For the Hermiticity-preserving maps, there are known bounds on the diamond norm [33, 38, 68]

$$\frac{1}{d}\|J(\Phi)\|_1 \leq \|\Phi\|_\diamond \leq \|\mathrm{Tr}_{\mathcal{Y}}|J(\Phi)|\|. \tag{2.52}$$

The diamond norm distance between quantum channels Φ_0 and Φ_1 is defined as $\|\Phi_0 - \Phi_1\|_\diamond$. The operational interpretation is as follows. We want to discriminate outputs of the channels using the Holevo-Helstrom theorem 1. We can optimize over the input states which can be entangled with an additional register. This operational interpretation for discrimination of quantum channels will be further discussed in Section 3.2.

2.3 Numerical range and support

2.3.1 Numerical range

In the studies on the symmetric discrimination of quantum measurements the key tool in the proofs will be the notion of numerical range. For a matrix $X \in \mathcal{L}(\mathcal{X})$, its *numerical range* is a subset of complex plane defined as

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

The essential property of the numerical range is known as the Hausdorf-Toeplitz theorem [69, 70], which states that $W(X)$ is a convex set. Hence it can be rewritten as

$$W(X) = \{\mathrm{Tr}(X\rho) : \ \rho \in \mathcal{D}(\mathcal{X})\}. \tag{2.54}$$

There are many generalizations of numerical range known in the literature [71–77] and one of them, which will be used a lot in this dissertation, is the q-numerical range [78–80]. The *q-numerical range* of the matrix X is defined as

$$W_q(X) \coloneqq \{\langle\varphi|X|\psi\rangle : \ \||\varphi\rangle\| = \||\psi\rangle\| = 1,\ \langle\phi|\psi\rangle = q,\ q \in \mathbb{C}\}. \tag{2.55}$$

Note that for $q = 1$ we recover the numerical range, that is $W_1(X) = W(X)$. Moreover, we will use the notation

$$\nu_q(X) \coloneqq \min\{|x| : \ x \in W_q(X)\}, \tag{2.56}$$

to denote the distance on the complex plane from the origin of the coordinate system to the q-numerical range. For simplicity, we will write $\nu(X)$ instead of $\nu_1(X)$.

The set $W_q(X)$ is compact and convex [78]. An important property of q-

numerical range is [81]

$$W_{q'} \subseteq \frac{q'}{q} W_q \quad \text{for} \quad q \leq q',\ q, q' \in \mathbb{R}. \tag{2.57}$$

Moreover, for every $q \in \mathbb{R}$ it holds that [81]

$$W_q\left(X \otimes \mathbb{1}\right) = W_q(X) \tag{2.58}$$

and therefore also

$$\nu_q\left(X \otimes \mathbb{1}\right) = \nu_q(X). \tag{2.59}$$

The detailed shape of q-numerical range is described in [79].

2.3.2 Supports

The notions of supports of quantum states and channels will be useful in numerous cases in this dissertation. Let $\rho \in \mathcal{D}(\mathcal{X})$ be a quantum state with spectral decomposition $\rho = \sum_{i=1}^{d} p_i |x_i\rangle\langle x_i|$. Then, the *support* of ρ is a subspace of $\mathcal{X}$ defined as

$$\operatorname{supp}(\rho) \coloneqq \operatorname{span}\{|x_i\rangle : p_i > 0\}. \tag{2.60}$$

Let $\Phi \in \mathcal{C}(\mathcal{X}, \mathcal{Y})$ be a quantum channel having Kraus operators $\{E_1, \ldots, E_k\} \subseteq \mathcal{L}(\mathcal{X}, \mathcal{Y})$. The *support* of the channel Φ is a subspace of $\mathcal{L}(\mathcal{X}, \mathcal{Y})$ which was defined in [82] as

$$\operatorname{supp}(\Phi) \coloneqq \operatorname{span}\{E_1, \ldots, E_k\}. \tag{2.61}$$

Chapter 3

Symmetric discrimination

In the symmetric discrimination scheme, we are given a black box which contains one of two quantum objects. Our goal will be to decide which of the objects was inside the black box and maximize the probability that this decision is correct. Discrimination of quantum channels and measurements is more complex than discrimination of quantum states. Quantum channels transform quantum states into quantum states. Quantum measurements take as input a quantum state and output a classical label. In both cases, we need to prepare some input state to get the output and gain knowledge about the content of the black box.

In this chapter, we will focus on the case when the black box can be used exactly once, and this will be called *single-shot* discrimination. We will be interested mostly in the discrimination of quantum measurements, so assume that the black box contains either the measurement $\mathcal{P}_0$ or the measurement $\mathcal{P}_1$. The classical descriptions of both measurements are known, so we can use this knowledge to prepare the discrimination strategy.

Before describing the strategy for the discrimination of quantum measurements and channels, let us quickly review the discrimination of quantum states [19–25]. Assume there are two quantum states and we know their classical descriptions. One of these states is secretly chosen with equal a priori probabilities and hidden inside the black box. To perform the discrimination, we can only measure the state in the black box. Such a measurement will be called a *final measurement* and denoted $\mathcal{P}_F$. The form of final measurement in the symmetric discrimination scheme is given by the Holevo-Helstrom Theorem 1.

In the most naïve approach to discrimination of quantum measurements, we can prepare some input state and measure it with the measurement contained in the black box. This scheme reduces to the problem of discrimination of probability distributions. However, not every input state will be equally good. The fundamental difficulty in such a naïve approach is finding the optimal input state to maximize the probability of correct discrimination. This scheme of discrimination will be

studied in Section 3.1, which will be based on [33].

A more sophisticated scheme for discrimination of quantum measurements allows for the use of entanglement [26, 83–88]. In this scheme, the measurement contained in the black box will act only on a part of an entangled input state. In other words, we can prepare as the input some entangled quantum state and perform the measurement contained in the black box on one part of this state. After this measurement, we obtain a classical label, and later, we can measure the remaining part of the entangled state by the final measurement $\mathcal{P}_F$. This is the most general scheme of discrimination of quantum measurements when the measurement in the black box can be used only once. It will be described in greater detail in Section 3.2.

Later, in Section 3.3, we will focus on the discrimination of von Neumann measurements. We will prove a condition when such measurements can be discriminated perfectly (with probability one). In the case when perfect discrimination will not be achievable, we will provide the optimal probability of their correct discrimination. This section will be based on the work [33]. The following section 3.4 will focus on the discrimination of SIC POVMs. We will calculate the probability of their correct discrimination and study when the use of entanglement improves the discrimination. This section contains results that have not been published.

3.1 Naïve discrimination

When a quantum state is measured, it ceases to exist. The only information given by the measurement is a classical label. Assume that in the black box there is one of two measurements which classical descriptions we know - either $\mathcal{P}_0$ or $\mathcal{P}_1$. In the most naïve scheme, to discriminate these measurements we only prepare an input state $|\psi\rangle$ and apply the measurement contained in the black box on the input state. As a result, we obtain a classical label, i, basing on which we perform some post-processing to decide whether the measurement was $\mathcal{P}_0$ or $\mathcal{P}_1$. This scheme of discrimination is depicted in Figure 3.1. The post-processing is denoted as a trivial final measurement $\mathcal{P}_F$. It allows us to decide, based on the label i, whether the final answer is 0, if the measurement was $\mathcal{P}_0$, or the final decision is 1, if the measurement in the back box was $\mathcal{P}_1$.

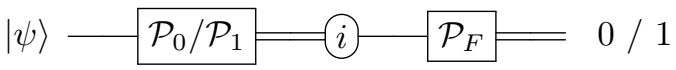

Figure 3.1: Scheme of naïve discrimination of quantum measurements $\mathcal{P}_0$ and $\mathcal{P}_1$.

Every quantum measurement can be seen as a quantum channel which outputs a probability distribution. More precisely, this quantum channel transforms the

input state into the diagonal matrix, where the elements on the diagonal correspond to probabilities of obtaining respective measurement labels. So does this problem of discrimination of quantum measurements directly reduce to discrimination of probability distributions? Although we can use the tools for discrimination of probability distributions, the essential part of the problem is choosing the optimal input state.

Finding the best input state is not trivial. Nevertheless, we still can calculate the probability of discrimination between von Neumann measurements without stating directly the form of the input state. This is formulated as the following Proposition, which was proved in [33].

Proposition 1 *Let $\mathcal{P}_0, \mathcal{P}_1$ be quantum measurements with effects $\{E_i\}_{i=1}^m$ and $\{F_i\}_{i=1}^m$ respectively. The probability p of their correct discrimination, without the usage of entangled states, is upper-bounded by*

$$p \leq \frac{1}{2} + \frac{1}{2} \max_{\Delta \subseteq \{1,\ldots,m\}} \left\| \sum_{i \in \Delta} (E_i - F_i) \right\|. \tag{3.1}$$

Proof. We will take advantage of the Holevo-Helstrom bound for the discrimination of quantum states (see Eq. (2.48)), which yields that the probability of correct discrimination between quantum states ρ_0 and ρ_1 is upper-bounded by $p \leq \frac{1}{2} + \frac{1}{4}\|\rho_0 - \rho_1\|_1$. In the scheme of discrimination of quantum measurements we can optimize over input states, and therefore we have the bound

$$p \leq \frac{1}{2} + \frac{1}{4} \max_{\rho} \|\mathcal{P}_0(\rho) - \mathcal{P}_1(\rho)\|_1. \tag{3.2}$$

To complete the proof we calculate

$$\begin{aligned}
\max_{\rho} \|\mathcal{P}_0(\rho) - \mathcal{P}_1(\rho)\|_1 &= \max_{\rho} \|\mathrm{diag}\left[(\mathcal{P}_0 - \mathcal{P}_1)(\rho)\right]\|_1 \\
&= \max_{\rho} \sum_i \left|\mathrm{Tr}\left(\rho(E_i - F_i)\right)\right| \\
&= \max_{|\psi\rangle} \sum_i \left|\langle\psi| \left(E_i - F_i\right) |\psi\rangle\right| \\
&= 2 \max_{\Delta \subseteq \{1,\ldots,m\}} \max_{|\psi\rangle} \langle\psi| \left(\sum_{i \in \Delta} (E_i - F_i)\right) |\psi\rangle \\
&= 2 \max_{\Delta \subseteq \{1,\ldots,m\}} \left\| \sum_{i \in \Delta} (E_i - F_i) \right\|.
\end{aligned} \tag{3.3}$$

where the third equality follows from the fact that a convex function achieves its

maximum on the boundary of its domain, which in our case is a pure state. ■

So far we were considering only the naïve scheme, where we only measure the prepared input state. But can we do better than this? The use of additional register often significantly improves the discrimination. This more general discrimination scheme will be discussed in the following section.

3.2 Entanglement-assisted discrimination

In this section, we will introduce the entanglement-assisted scheme of discrimination of quantum channels and measurements. First, we will describe this scheme for quantum channels and later, we will see how to modify it to discriminate quantum measurements.

The scheme of discrimination of quantum channels Φ_0 and Φ_1 is depicted in Figure 3.2.

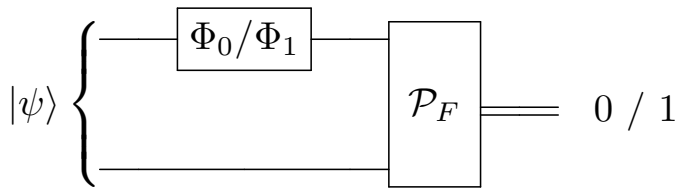

Figure 3.2: Scheme of single-shot discrimination of quantum channels Φ_0 and Φ_1.

We begin with preparing an entangled input state $|\psi\rangle$ on two registers. Then, on the first register we apply the channel is contained in the black box – either Φ_0 or Φ_1. Then, we prepare a final measurement $\mathcal{P}_F$ and measure both systems. Basing on the outcome of the final measurement we make a decision which of the two channels was inside the black box. If the label of the final measurement was 0, then we say that in the black box was Φ_0. Similarly, when the label of the final measurement was 1, then we say that in the black box was Φ_1.

The probability of successful discrimination between these channels can be calculated from the Holevo-Helstrom theorem [66, 89]. This theorem was stated for the discrimination of quantum states in Eq. (2.48). Here, we will be interested in another version of this theorem, which works for the discrimination of quantum channels, More precisely, assuming that one of the channels, either Φ_0 or Φ_1, is given with equal probability, the bound on the probability of correct discrimination between these channels is upper bounded by

$$p \leq \frac{1}{2} + \frac{1}{4}\|\Phi_0 - \Phi_1\|_\diamond, \tag{3.4}$$

where the distance between Φ_0 and Φ_1 is expressed in terms of the diamond norm defined in Eq. (2.50)

Now we proceed to discrimination of quantum measurements. The scheme of discrimination of quantum measurements $\mathcal{P}_0$ and $\mathcal{P}_1$ is presented in Figure 3.3.

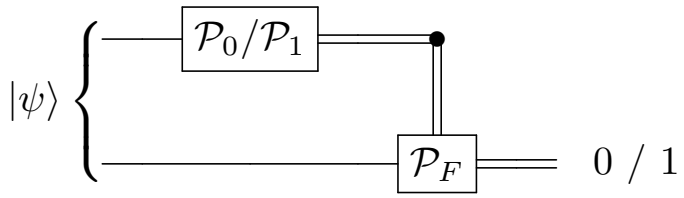

Figure 3.3: Scheme of single-shot discrimination of quantum measurements $\mathcal{P}_0$ and $\mathcal{P}_1$.

The first step of the discrimination scheme is preparing an input state $|\psi\rangle$ on the compound register. Naturally, we can choose an entangled input state. Then, we measure the first register by the measurement contained in the black box – either $\mathcal{P}_0$ or $\mathcal{P}_1$, and obtain a classical label. Basing on this label we prepare a final measurement $\mathcal{P}_F$ and measure the second register. Eventually, we make a decision whether in the black box there was $\mathcal{P}_0$ or $\mathcal{P}_1$.

The schemes of discrimination of quantum channels and measurements have many things in common. In both cases we prepare an entangled input state and apply the black box on only one part of this state. Moreover, in both cases we make a decision about the content of the black box basing on the outcome of the final measurements. The key difference is the application of the final measurement. For discrimination of quantum channels, the final measurement is used to measure both registers. On the other hand, when discriminating quantum measurements, after applying the measurement in the black box we have a classical label. Basing on this label we prepare a final measurement and measure only the second register.

How to calculate the probability of successful discrimination of quantum measurements? We can use the Holevo-Helstrom theorem for discrimination of quantum channels (see Eq. (3.4)). To do this, we need to know that every quantum measurement $\mathcal{P}$ with effects $\{E_1, \ldots, E_m\}$ can be associated with a quantum channel, which action on a quantum state ρ can be expressed as [38].

$$\mathcal{P}(\rho) = \sum_{i=1}^{m} \mathrm{Tr}(E_i \rho) |i\rangle\langle i|. \tag{3.5}$$

This is a quantum-classical channel which output is diagonal, and i-th element on the diagonal corresponds to the probability of obtaining i-th measurement label. As we will be mostly interested in discrimination of measurements with rank-one

effects $\{|x_i\rangle\langle x_i|\}_i$, the above formula can be written as

$$\mathcal{P}(\rho) = \sum_{i=1}^{m} \langle x_i|\rho|x_i\rangle |i\rangle\langle i|. \tag{3.6}$$

Thanks to this representation, from the Holevo-Helstrom we know that the upper bound on the probability of correct discrimination between quantum measurements $\mathcal{P}_0$ and $\mathcal{P}_1$ yields

$$p \leq \frac{1}{2} + \frac{1}{4}\|\mathcal{P}_0 - \mathcal{P}_1\|_\diamond. \tag{3.7}$$

The strategy for discrimination of quantum measurements and channels include preparing input state and final measurement. Let us emphasize here that the form of the final measurement is given by the Holevo-Helstrom theorem, so we do not need to search for the best measurement. Nevertheless, the input state needs to be optimized and there is no general rule for it.

Later in this dissertation, we will be mostly interested in the entanglement-assisted discrimination scheme. Therefore, whenever we will be talking about a discrimination scheme, we will implicitly assume that the use of additional system and entangled input is allowed.

3.3 Discrimination of von Neumann measurements

In this section we will be interested in (entanglement-assisted) discrimination between two von Neumann measurements. Without loss of generality, for the symmetric discrimination we can assume that one of the measurements is in the canonical basis [33, 35]. In other words, we are given a black box which contains either $\mathcal{P}_0 \coloneqq \mathcal{P}_U$, where U is a fixed unitary matrix, or $\mathcal{P}_1 \coloneqq \mathcal{P}_{\mathbf{1}}$. We will use the entanglement-assisted scheme introduced in the previous section to decide which of the measurements was hidden the black box.

As von Neumann measurements can be parametrized by unitary matrices, it should come as no surprise that discrimination of von Neumann measurements is closely related with the task of discrimination of unitary channels [90–96]. Therefore, we will often use the well-known [38] result for the diamond norm distance between unitary channels. This result relates the problem of discrimination of unitary channels with the notion of numerical range defined in Eq. (2.53).

Proposition 2 ([38]) *Let $U \in \mathcal{U}(\mathcal{X})$ and $\Phi_U \in \mathcal{C}(\mathcal{X})$ be a unitary channel. Let $\Phi_{\mathbf{1}} = \mathbb{1}_{\mathcal{L}(\mathcal{X})}$ be an identity channel. Then*

$$\|\Phi_U - \Phi_{\mathbf{1}}\|_\diamond = 2\sqrt{1-\nu^2}, \tag{3.8}$$

where $\nu = \min\{|x| : x \in W(U^\dagger)\}$.

The rest of this section is organized as follows. We will begin with the case of perfect discrimination, that is we will state a condition when quantum measurement $\mathcal{P}_U$ can be discriminated from $\mathcal{P}_{\mathbb{1}}$ with probability one after a single query to the black box. Later, we will consider more general case when perfect discrimination may not be achieved after just one query. The main part of this section will be a theorem stating the probability of correct discrimination in the single-shot case [33]. We will also present its geometrical representation.

The following proposition gives a condition when two von Neumann measurements can be discriminated perfectly in the single-shot scenario. This proposition was first proved in [97], and later, independently, in [33].

Proposition 3 *Let* $U \in \mathcal{U}(\mathcal{X})$. *Von Neumann measurements* $\mathcal{P}_U$ *and* $\mathcal{P}_{\mathbb{1}}$ *can be discriminated perfectly in the single-shot scenario if and only if there exists a state* $\rho \in \mathcal{D}(\mathcal{X})$ *such that*

$$\operatorname{diag}(U^\dagger \rho) = 0. \tag{3.9}$$

Proof. Let $\rho \in \mathcal{D}(\mathcal{X})$ be a quantum state satisfying the alternative formula for the diamond norm in Eq. (2.51), that is

$$\|\mathcal{P}_U - \mathcal{P}_{\mathbb{1}}\|_\diamond = \|(\mathbb{1} \otimes \sqrt{\rho}) J(\mathcal{P}_U - \mathcal{P}_{\mathbb{1}})(\mathbb{1} \otimes \sqrt{\rho})\|_1 . \tag{3.10}$$

We calculate the diamond norm distance between the POVMs

$$\begin{aligned}
\|\mathcal{P}_U - \mathcal{P}_{\mathbb{1}}\|_\diamond &= \left\| (\mathbb{1} \otimes \sqrt{\rho}) \left(\sum_{i=1}^{d} |i\rangle\langle i| \otimes (|u_i\rangle\langle u_i| - |i\rangle\langle i|)^\top \right) (\mathbb{1} \otimes \sqrt{\rho}) \right\|_1 \\
&= \left\| \sum_{i=1}^{d} |i\rangle\langle i| \otimes (\sqrt{\rho}\,(|u_i\rangle\langle u_i| - |i\rangle\langle i|)\,\sqrt{\rho}) \right\|_1 \\
&= \sum_{i=1}^{d} \operatorname{tr}\left| \sqrt{\rho}|u_i\rangle\langle u_i|\sqrt{\rho} - \sqrt{\rho}|i\rangle\langle i|\sqrt{\rho} \right| \\
&= \sum_{i=1}^{d} \operatorname{tr}\left| \langle u_i|\rho|u_i\rangle \frac{\sqrt{\rho}|u_i\rangle}{\sqrt{\langle u_i|\rho|u_i\rangle}} \frac{\langle u_i|\sqrt{\rho}}{\sqrt{\langle u_i|\rho|u_i\rangle}} - \langle i|\rho|i\rangle \frac{\sqrt{\rho}|i\rangle}{\sqrt{\langle i|\rho|i\rangle}} \frac{\langle i|\sqrt{\rho}}{\sqrt{\langle i|\rho|i\rangle}} \right| \\
&= \sum_{i=1}^{d} \sqrt{\left(\langle u_i|\rho|u_i\rangle + \langle i|\rho|i\rangle\right)^2 - 4\,|\langle u_i|\rho|i\rangle|^2},
\end{aligned} \tag{3.11}$$

where the last equality is a direct application of Lemma 1.

Now we proceed to proving the direct implication. Assume that $\mathcal{P}_U$ and $\mathcal{P}_{\mathbf{1}}$ can be discriminated perfectly, that is $\|\mathcal{P}_U - \mathcal{P}_{\mathbf{1}}\|_\diamond = 2$. Suppose by contradiction that the condition in Eq (3.9) is not satisfied. Therefore, for every state ρ there exists i such that $\langle u_i|\rho|i\rangle \neq 0$. Hence we have the inequality

$$\sum_{i=1}^{d} \sqrt{\left(\langle u_i|\rho|u_i\rangle + \langle i|\rho|i\rangle\right)^2 - 4\left|\langle u_i|\rho|i\rangle\right|^2} < \sum_{i=1}^{d} \left(\langle u_i|\rho|u_i\rangle + \langle i|\rho|i\rangle\right) = 2. \quad (3.12)$$

Therefore $\|\mathcal{P}_U - \mathcal{P}_{\mathbf{1}}\|_\diamond < 2$, which gives a contradiction.

As for the reverse implication, assume there exists a state ρ satisfying Eq. (3.9), that is, $\langle u_i|\rho|i\rangle = 0$ for every i. From Eq. (3.11) we obtain $\|\mathcal{P}_U - \mathcal{P}_{\mathbf{1}}\|_\diamond = 2$, which means that the measurements $\mathcal{P}_U$ and $\mathcal{P}_{\mathbf{1}}$ can be discriminated perfectly in the single-shot scenario. ■

When the condition in the above proposition is not fulfilled, then the single-shot scenario does not allow for perfect discrimination. In such case, we still can try to discriminate the measurements minimizing the probability of making a mistake. So what is the optimal probability of successful discrimination we can achieve for a given pair of von Neumann measurements? Here we arrive at the main theorem of this section. This theorem gives a formula for the diamond norm distance between any two von Neumann measurements. This theorem was proved in [33], and its proof requires a great deal of additional technical results, which are presented in the Appendix A.

Theorem 2 *Let $U \in \mathcal{U}(\mathcal{X})$ and let $\mathcal{P}_U$ and $\mathcal{P}_{\mathbf{1}}$ be von Neumann measurements. Let $\mathcal{DU}(\mathcal{X})$ be the set of diagonal unitary matrices of dimension d and Φ_U be a unitary channel. Then*

$$\|\mathcal{P}_U - \mathcal{P}_{\mathbf{1}}\|_\diamond = \min_{E \in \mathcal{DU}(\mathcal{X})} \|\Phi_{UE} - \Phi_{\mathbf{1}}\|_\diamond. \quad (3.13)$$

The probability of correct discrimination between two von Neumann measurements in the single-shot scenario can be calculated directly from the Holevo-Helstrom theorem (Eq. (3.4)). This is formulated as the following corollary.

Corollary 1 *With the notation as in Theorem 2, the probability of correct discrimination between $\mathcal{P}_U$ and $\mathcal{P}_{\mathbf{1}}$ in the single-shot scenario is upper bounded as follows*

$$p \leq \frac{1}{2} + \frac{1}{4} \min_{E \in \mathcal{DU}(\mathcal{X})} \|\Phi_{UE} - \Phi_{\mathbf{1}}\|_\diamond. \quad (3.14)$$

Let us finish this section with the geometrical interpretation of the diamond norm distance between von Neumann measurements. For a given unitary matrix $U \in \mathcal{U}(\mathcal{X})$ of dimension d, we define $\Theta(U)$ as the angle of the shortest arc containing all eigenvalues of U. Let us define an optimized version of Θ as

$$\Upsilon(U) \coloneqq \min_{E \in \mathcal{DU}(\mathcal{X})} \Theta(UE). \tag{3.15}$$

Note that in this definition there is the same minimization as in Theorem 2.

Knowing that the distance between unitary channels is expressed by the distance from zero to the numerical range of a unitary matrix (see Prop. 2), we can generalize this interpretation to the case of von Neumann measurements. It is presented in Figure 3.4.

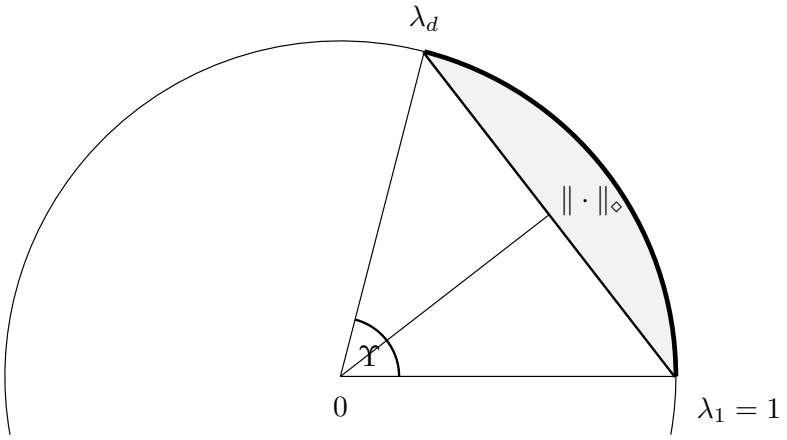

Figure 3.4: Geometrical interpretation of the diamond norm distance between von Neumann measurements $\mathcal{P}_U$ and $\mathcal{P}_{\mathbb{1}}$. λ_1 and λ_d denote the most distant eigenvalues of a unitary matrix UE_0, where E_0 is the optimal diagonal unitary matrix in Eq. (3.15). The numerical range of UE_0 is contained in the gray area.

This figure presents a sector of the complex place with a part of the unit circle. Let λ_1 and λ_d be the most distant eigenvalues of UE_0, where E_0 is the optimal diagonal unitary matrix in Eq. (3.15). In other words, all the other eigenvalues $\lambda_2, \ldots, \lambda_{d-1}$ lie on the unit circle between λ_1 and λ_d. Without loss of generality we can take $\lambda_1 = 1$. The numerical range of the unitary matrix UE_0 is the polygon connecting all its eigenvalues. It the picture it is contained in the gray area. The diamond norm distance between $\mathcal{P}_U$ and $\mathcal{P}_{\mathbb{1}}$ corresponds to the distance between λ_1 and λ_d.

3.4 Discrimination of SIC POVMs

So far, we were studying discrimination of von Neumann measurements, which effects are rank-one projection onto orthogonal subspaces. Hence, every von Neumann measurement of dimension d has exactly d effects. In the remaining part of this chapter we will be interested in discrimination of more general measurements with rank-one effects. We will write *rank-one measurements* when talking about measurements with rank-one effects.

Let us begin with considering the number of effects. How many effects at most can a rank-one measurement have? The minimum number is the same as the dimension of the measurement, which is the case of von Neumann measurements. Nevertheless, the rank-one measurement of dimension d can have more then d effects. The maximal number of linearly independent effects is d^2.

Unfortunately, a rank-one measurement which has more than d effects cannot be simply parametrized by a unitary matrix, so we cannot use the results known for discrimination of unitary channels. Fortunately, there is a class of rank-one measurements having very useful symmetry properties. These measurements are known as symmetric informationally complete (SIC) POVMs. Let us recall that a SIC POVM $\mathcal{P}$ of dimension d has d^2 effects $\{E_1, \ldots, E_{d^2}\}$, where $E_i = \frac{1}{d}|\phi_i\rangle\langle\phi_i|$ and $\||\phi_i\rangle\| = 1$ for every $i = 1, \ldots, d^2$. Moreover, the symmetry condition states that

$$|\langle\phi_i|\phi_j\rangle|^2 = \frac{1}{d+1} \tag{3.16}$$

for $i \neq j$.

In this section we will study discrimination of SIC POVMs $\mathcal{P}_0$ and $\mathcal{P}_1$ of dimension d having effects $\{E_1, \ldots, E_{d^2}\}$ and $\{F_1, \ldots, F_{d^2}\}$, respectively. We will assume that effects of both measurements $\mathcal{P}_0$ and $\mathcal{P}_1$ are related by a permutation. More specifically, for a permutation π of d^2 elements we will assume that $F_i = E_{\pi(i)}$.

Recall from Subsection 2.1.3 in the Preliminaries, that is is an open question whether SIC POVMs exist in every dimension [62]. In this section, we will simply assume that they do exist for the studied dimensions. We will explore how SIC POVMs can be discriminated without going into details of properties of specific dimensions. Hence, in the formulations of theorems concerning SIC POVMs of bigger dimensions, we will omit the assumption that the SIC POVMs in these dimensions do exist.

From Holevo-Helstrom theorem it holds that the probability of correct discrimination between SIC POVMs $\mathcal{P}_0$ and $\mathcal{P}_1$ is upper-bounded by $p \leq \frac{1}{2} + \frac{1}{4}\|\mathcal{P}_0 - \mathcal{P}_1\|_\diamond$. In this section we will first state a proposition which gives lower and upper bounds on the diamond norm distance between two SIC POVMs. Later, we will provide some conclusions which can be drawn from this proposition. Finally, we will study single-shot discrimination of SIC POVMs when their dimension tends to

infinity. The results presented in this section have not been published before this dissertation.

Proposition 4 *Let π be a permutation of d^2 numbers with k fixed points. Let $\mathcal{P}_0 = \{E_1, \ldots, E_{d^2}\}$ and $\mathcal{P}_1 = \{F_1, \ldots, F_{d^2}\}$, where $F_i = E_{\pi(i)}$ for every $i = 1, \ldots, d^2$, be SIC POVMs. Then*

$$\frac{2\left(d^2-k\right)}{\sqrt{d^3\left(d+1\right)}} \leq \left\|\mathcal{P}_0 - \mathcal{P}_1\right\|_\diamond \leq \frac{d^2-k}{\sqrt{d(d+1)}}. \tag{3.17}$$

Proof. We will use the following well-known bounds on the diamond norm from Eq. (2.52)

$$\frac{1}{d}\|J(\Phi)\|_1 \leq \|\Phi\|_\diamond \leq \|\mathrm{Tr}_1|J(\Phi)|\|. \tag{3.18}$$

As for the lower bound, from the properties of the trace norm we have

$$\frac{1}{d}\left\|J(\mathcal{P}_0) - J(\mathcal{P}_1)\right\|_1 = \frac{1}{d}\left\|\sum_{i=1}^{d^2} |i\rangle\langle i| \otimes \left(E_i - F_i\right)^\top\right\|_1 = \frac{1}{d}\sum_{i=1}^{d^2}\left\|E_i - F_i\right\|_1, \tag{3.19}$$

and using Lemma 1 for $\alpha = \beta = \frac{1}{d}$ and $|\langle u|v\rangle|^2 = \frac{1}{d+1}$ we obtain

$$\left\|E_i - F_i\right\|_1 = \sqrt{\left(\frac{2}{d}\right)^2 - \frac{4}{d^2}\frac{1}{d+1}} = \sqrt{\frac{4}{d^2}\left(1 - \frac{1}{d+1}\right)} = \frac{2}{\sqrt{d\left(d+1\right)}}. \tag{3.20}$$

Thus, the lower bound can be expressed as

$$\frac{1}{d}\left\|J(\mathcal{P}_0) - J(\mathcal{P}_1)\right\|_1 = \frac{1}{d}\left(d^2-k\right)\frac{2}{\sqrt{d\left(d+1\right)}} = \frac{2\left(d^2-k\right)}{\sqrt{d^3\left(d+1\right)}}. \tag{3.21}$$

Now we consider the upper bound. First, from the properties of the spectral norm we have

$$\begin{aligned}\left\|\mathrm{Tr}_1\left|J(\mathcal{P}_0) - J(\mathcal{P}_1)\right|\right\|_\infty &= \left\|\mathrm{Tr}_1\left|\sum_{i=1}^{d^2} |i\rangle\langle i| \otimes \left|E_i - F_i\right|^\top\right|\right\| \\ &= \left\|\sum_{i=1}^{d^2}\mathrm{Tr}_1\left(|i\rangle\langle i| \otimes \left|E_i - F_i\right|^\top\right)\right\| = \left\|\sum_{i=1}^{d^2}\left|E_i - F_i\right|\right\|,\end{aligned} \tag{3.22}$$

and from the triangle inequality and the expression for the greatest eigenvalue from

Lemma 1 we obtain

$$\left\| \sum_{i=1}^{d^2} |E_i - F_i| \right\| \leq \sum_{i=1}^{d^2} \| |E_i - F_i| \| = \frac{d^2 - k}{\sqrt{d(d+1)}}. \tag{3.23}$$

■

The bounds in Proposition 4 depend only on the dimension of the measurement, d, and the number of fixed points, k, of the permutation π. A simple observations is that if the permutation π has sufficiently many fixed points, the measurements $\mathcal{P}_0$ and $\mathcal{P}_1$ will not be discriminated perfectly. More precisely, if

$$k \geq \left\lceil d^2 - 2\sqrt{d(d+1)} \right\rceil, \tag{3.24}$$

then the upper bound for the diamond norm is smaller than two, and thus, these two SIC POVMs cannot be discriminated perfectly in the single-shot scenario.

A few additional facts about discrimination of qubit SIC POVMs can be extracted from Proposition 4, which are formulated as two complementary corollaries. Corollary 2 provides a condition when the use of entangled input state does not improve the discrimination. Corollary 3 focuses on the qubit case. It states the exact value of the diamond norm distance between $\mathcal{P}_0$ and $\mathcal{P}_1$ and points when the use of entangled input state indeed improves the discrimination.

Corollary 2 *Let $\mathcal{P}_0 = \{E_1, \ldots, E_{d^2}\}$ and $\mathcal{P}_1 = \{F_1, \ldots, F_{d^2}\}$, where $F_i = E_{\pi(i)}$ for every $i = 1, \ldots d^2$, be SIC POVMs. Let π be the permutation with $d^2 - 2$ fixed points. Then the use of entanglement does not give any advantage in discrimination between $\mathcal{P}_0$ and $\mathcal{P}_1$.*

Proof. The upper bound for the discrimination without entanglement can be calculated by the use of Proposition 1 as

$$p \leq \frac{1}{2} + \frac{1}{4} \max_{\rho} \|\mathrm{diag}[(\mathcal{P}_0 - \mathcal{P}_1)(\rho)]\|_1 = \frac{1}{2} + \frac{1}{2} \max_{\Delta \subseteq \{1,\ldots,d^2\}} \left\| \sum_{i \in \Delta} (E_i - F_i) \right\|. \tag{3.25}$$

Using Lemma 1 we calculate

$$\begin{aligned} \max_{\rho} \|\mathrm{diag}[(\mathcal{P}_0 - \mathcal{P}_1)(\rho)]\|_1 &= 2 \max_{\Delta \subseteq \{1,\ldots,d^2\}} \left\| \sum_{i \in \Delta} (E_i - F_i) \right\| \\ = 2 \max_i \|E_i - F_i\| &= 2\sqrt{\left(\frac{2}{d}\right)^2 - \frac{4}{d^2}\frac{1}{d+1}} = \frac{2}{\sqrt{d(d+1)}}. \end{aligned} \tag{3.26}$$

The upper bound for the diamond norm from Eq. (2.52) is given by

$$\|\mathrm{Tr}_1 |J(\mathcal{P}_0) - J(\mathcal{P}_1)|\| = \frac{d^2 - (d^2-2)}{\sqrt{d(d+1)}} = \frac{2}{\sqrt{d(d+1)}}, \tag{3.27}$$

and we have the equality. ■

Note that the above corollary provides only the information that the use of entanglement does not improve the discrimination when if the permutation π has $d^2 - 2$ fixed points. Although there is no general condition when the use of entanglement improves the single-shot discrimination, the following corollary fully characterizes the qubit case.

Corollary 3 *Let $\mathcal{P}_0 = \{E_1, E_2, E_3, E_4\}$ and $\mathcal{P}_1 = \{F_1, F_2, F_3, F_4\}$, where $F_i = E_{\pi(i)}$, be SIC POVMs. Then $\|\mathcal{P}_0 - \mathcal{P}_1\|_\diamond = \frac{4-k}{\sqrt{6}}$, where k is the number of fixed points of the permutation π. Moreover, if the permutation π has either zero or one fixed point, then the discrimination with the use of entanglement always outperforms the discrimination without the use of entanglement.*

Proof. First, basing on Proposition 4, we note that for $d = 2$ the lower bound on the diamond norm distance between $\mathcal{P}_0$ and $\mathcal{P}_1$ is equal to the upper bound and yields

$$\|\mathcal{P}_0 - \mathcal{P}_1\|_\diamond = \frac{4-k}{\sqrt{6}}. \tag{3.28}$$

In order to obtain the probability of correct discrimination without the use of entanglement from Proposition 1, we need to calculate the value

$$\max_\rho \|\mathcal{P}_0(\rho) - \mathcal{P}_1(\rho)\|_1 = 2 \max_{\Delta \subseteq \{1,\ldots,4\}} \left\| \sum_{i \in \Delta} (E_i - F_i) \right\|. \tag{3.29}$$

Consider first the case of a permutation with no fixed point. Then, the maximal value $\Delta \subseteq \{1, \ldots, 4\}$ is achieved for 2-element subset of Δ and

$$\max_\rho \|\mathcal{P}_0(\rho) - \mathcal{P}_1(\rho)\|_1 = 2 \|E_i + E_j - E_k - E_l\| = \frac{2\sqrt{3}}{3}. \tag{3.30}$$

If the permutation has one fixed point, then the calculations simplify to the use of Lemma 1 and calculating the largest eigenvalue of $\|E_i - E_j\|$. Eventually, for a permutation with no fixed point we have

$$\max_\rho \|\mathcal{P}_0(\rho) - \mathcal{P}_1(\rho)\|_1 = \frac{2\sqrt{3}}{3} < \frac{2\sqrt{6}}{3} = \|\mathcal{P}_0 - \mathcal{P}_1\|_\diamond \tag{3.31}$$

while for a permutation with one fixed point we have

$$\max_{\rho} \|\mathcal{P}_0(\rho) - \mathcal{P}_1(\rho)\|_1 = \frac{\sqrt{6}}{3} < \frac{\sqrt{6}}{2} = \|\mathcal{P}_0 - \mathcal{P}_1\|_\diamond, \tag{3.32}$$

thus the use of entanglement improves the discrimination. The case of a permutation with two fixed points follows from Corollary 2. ■

The above corollary characterized the two-dimensional case. Now we will study the discrimination for big dimensions. More precisely, we will see that in the asymptotic limit, when the dimension tends to infinity, the chances of successful discrimination are very big.

Theorem 3 *Let π_d be a uniformly chosen random permutation on d^2 elements. Assume that a SIC POVM of dimension d exists. Let $\mathcal{P}_0 = \{E_1, \ldots, E_{d^2}\}$ and $\mathcal{P}_1 = \{F_1, \ldots, F_{d^2}\}$ be SIC POVMs where $F_i = E_{\pi_d(i)}$ for some permutation π_d. Then the measurements $\mathcal{P}_0$ and $\mathcal{P}_1$ can be discriminated perfectly almost surely as d tends to infinity.*

Proof. Let X_d denote the random variable which takes the value of the lower bound on the diamond norm between a SIC POVM of dimension d and a SIC POVM permuted according to permutation π_d.

In order to prove that $X_d \xrightarrow{a.s.} 2$ it suffices to show that for any $\epsilon > 0$ we have that

$$\sum_{d=1}^{\infty} P\left(|X_d - 2| > \epsilon\right) < \infty. \tag{3.33}$$

Let $\star(\pi_d)$ be a random variable which denotes the number of fixed points of the permutation π_d. Then using the lower bound for the diamond norm from Proposition 4 we obtain

$$X_d(\pi_d) = \frac{2(d^2 - n)}{\sqrt{d^3(d+1)}} \mathbb{1}_{\{\star(\pi_d)=n\}}. \tag{3.34}$$

$$P\left(X_d = \frac{2(d^2 - n)}{\sqrt{d^3(d+1)}}\right) = P\left(\star(\pi_d) = n\right). \tag{3.35}$$

Now, if $\star(\pi_d) = n$, then

$$\begin{aligned} P(|X_d - 2| > \epsilon) &= P\left(\left|\frac{2(d^2 - \star(\pi_d))}{\sqrt{d^3(d+1)}} - 2\right| > \epsilon\right) = P\left(2 - \frac{2(d^2 - \star(\pi_d))}{\sqrt{d^3(d+1)}} > \epsilon\right) \\ &= P\left(\star(\pi_d) > d^2 - \frac{(2-\epsilon)\sqrt{d^3(d+1)}}{2}\right) \\ &\leq P\left(\star(\pi_d) > \frac{\epsilon}{2}d^2 - \frac{d}{2}\right), \end{aligned} \tag{3.36}$$

where the last inequality comes from the fact that

$$d^2 - \frac{(2-\epsilon)\sqrt{d^3(d+1)}}{2} \geq \frac{\epsilon}{2}d^2 - \frac{d}{2}. \tag{3.37}$$

Knowing that

$$P(\star(\pi_d) = n) = \frac{1}{n!}\sum_{i=0}^{d^2-n}\frac{(-1)^i}{i!} \tag{3.38}$$

we calculate

$$P(\star(\pi_d) > n) = \sum_{m=n+1}^{d^2}\frac{1}{m!}\sum_{i=0}^{d^2-m}\frac{(-1)^i}{i!} \leq \sum_{m=n+1}^{d^2}\frac{1}{m!}\sum_{i=0}^{\infty}\frac{1^i}{i!} = \mathrm{e}\sum_{m=n+1}^{d^2}\frac{1}{m!}. \tag{3.39}$$

Finally

$$\begin{aligned} \sum_{d=1}^{\infty} P(|X_d - 2| > \epsilon) &\leq \sum_{d=1}^{\infty} P\left(\star(\pi_d) > \frac{\epsilon}{2}d^2 - \frac{d}{2}\right) \leq \sum_{d=1}^{\infty}\mathrm{e}\sum_{m=\left\lfloor\frac{\epsilon}{2}d^2 - \frac{d}{2}+1\right\rfloor}^{d^2}\frac{1}{m!} \\ &\leq \mathrm{e}\sum_{d=1}^{\infty} d^2\frac{1}{\left\lfloor\frac{\epsilon}{2}d^2 - \frac{d}{2}\right\rfloor!} < \infty. \end{aligned} \tag{3.40}$$

and from the Borel-Cantelli lemma we obtain almost sure convergence. ∎

Chapter 4

Symmetric multiple-shot discrimination

In the previous chapter, we studied symmetric discrimination in the single-shot case. We assumed that the black box containing one of two quantum objects could be used only once. This assumption is a natural starting point for studying the discrimination of quantum objects. Nonetheless, such a discrimination scheme is not too general.

Before we describe the multiple-shot discrimination of quantum channels and measurements, let us quickly review this problem for the discrimination of quantum states. When a black box contains a quantum state, to get some information about the state, we need to measure it. However, measuring the state destroys it. As the output, we obtain only a classical label, and the quantum state ceases to exist. How would multiple-shot discrimination work in the case when the black box contained a quantum state? We can prepare many copies of the same black box and measure each of them. Nevertheless, if we cannot perfectly discriminate between quantum states in the single-shot scenario, then perfect discrimination cannot be achieved after any finite number of queries. [98, 99].

When the black box contains either a quantum channel or a measurement, nothing prevents us from using this black box many times. When we discriminate quantum channels as the output from the black box, we obtain a quantum state. This state can also be used as the input in the following query to the black box, and this procedure can be repeated many times. Such a discrimination scheme is known as a *sequential* scheme, but this is only one of many possible discrimination schemes. The sequential scheme is significantly useful for discrimination of unitary channels – it allows for perfect discrimination of unitary channels without the use of entangled states. [95]

When discriminating quantum measurements as the output, we obtain a classical label, and the measured quantum state ceases to exist. However, as it was described

in the entanglement-assisted scheme for discrimination of quantum measurements, we can prepare an input state on a compound space and measure only part of it. How to generalize this scheme? For instance, we can prepare an input state on more than two registers and apply the black box on more than one register. We can also use the classical label to prepare another input state for the following query to the black box. As we can see, there are many possibilities of generalizing the single-shot discrimination scheme [67, 100, 101]. In this chapter, we will focus on the parallel and adaptive schemes, which will be described in detail in Sections 4.2 and 4.3, respectively.

Later in this chapter, we will study the discrimination of von Neumann measurements in Section 4.4. This section will cover the results proved in [35]. We will calculate the probability of correct discrimination of such measurements after N queries in parallel. We will prove that for discrimination of von Neumann measurement, the use of an adaptive scheme does not give any advantage over the parallel one. Next, we will focus on the discrimination of SIC POVMs in Section 4.5, which will be cased on [34] We will see how an adaptive scheme can improve discrimination.

Before going into details of the parallel and adaptive discrimination schemes, let us elaborate on the primary goal of multiple-shot discrimination. Recall from the previous chapter that the term *perfect discrimination* refers to the identification of the content of the black box with probability one. We saw in the previous chapter that it is not always possible to achieve perfect discrimination in the single-shot scheme. When after a single query to the black box, we cannot discriminate its content perfectly, we can give the multiple-shot scheme a try. And here arises a question of whether we can always achieve perfect discrimination after some finite number of queries to the black box. Unfortunately, this is not the case. Conditions when perfect discrimination can be achieved in a finite number of queries, will be studied in Section 4.1.

4.1 Conditions for perfect discrimination

The situation of special importance happens when we can be sure that the discrimination procedure gives the correct answer. In the single-shot case, quantum channels can be discriminated *perfectly*, when the right-hand side of Eq. (3.4) equals one, that is when $\|\Phi_0 - \Phi_1\|_\diamond = 2$. Nevertheless, even if the quantum channels cannot be discriminated perfectly in the single-shot case, it does not mean that they cannot be discriminated perfectly at all, as it would be for the discrimination of quantum states. Recall that when quantum states cannot be discriminated perfectly after one query, they cannot be discriminated perfectly after any finite number of queries. [98, 99] In the work [81], the authors derived a condition when quantum

channels can be discriminated perfectly after a finite number of queries to the black box. Their condition is very general and does not specify what discrimination scheme should be used.

Before stating this condition, let us define the notion of *disjointness* of quantum states and channels. Quantum states ρ_0 and ρ_1 are *disjoint* when $\mathrm{supp}(\rho_0) \cap \mathrm{supp}(\rho_1) = \{0\}$. The notion of disjointness can be generalized to quantum channels. We have the following definition.

Definition 3 *Two quantum channels $\Phi_0, \Phi_1 \in \mathcal{C}(\mathcal{X})$ are called (entanglement-assisted) disjoint if there exists an input state $|\psi\rangle \in \mathcal{D}(\mathcal{X} \otimes \mathcal{X})$ such that the output states $\left(\Phi_0 \otimes \mathbb{1}_{\mathcal{L}(\mathcal{X})}\right)(|\psi\rangle\langle\psi|)$ and $\left(\Phi_1 \otimes \mathbb{1}_{\mathcal{L}(\mathcal{X})}\right)(|\psi\rangle\langle\psi|)$ are disjoint.*

The condition when quantum channels can be discriminated perfectly in a finite number of queries is formulated as the following theorem.

Theorem 4 ([81]) *Let Φ_0 and Φ_1 be quantum channels having Kraus operators $\{M_i\}_i$ and $\{N_j\}_j$ respectively. Quantum channels Φ_0 and Φ_1 are can be discriminated perfectly by a finite number of queries if and only if they are (entanglement-assisted) disjoint and $\mathbb{1} \notin \mathrm{span}\{M_i^\dagger N_j\}_{i,j}$.*

Perfect discrimination of quantum measurements is not always achievable. As quantum measurements can be seen as quantum channels which give diagonal outputs, we can use the Theorem 4 to check if a given pair of quantum measurements can be discriminated perfectly after a finite number of queries to the black box. We will be mostly interested in discrimination of measurements with rank-one effects, so let us formulate a corollary, when a pair of such rank-one measurements can be discriminated perfectly after a finite number of queries.

Corollary 4 *Let $\mathcal{P}_0$ and $\mathcal{P}_1$ be POVMs of dimension d with effects $\{|x_i\rangle\langle x_i|\}_{i=1}^m$ and $\{|y_i\rangle\langle y_i|\}_{i=1}^m$ respectively. Then $\mathcal{P}_0$ and $\mathcal{P}_1$ are can be discriminated perfectly after a finite number of uses if and only if*

- $\mathbb{1} \notin \mathrm{span}\{|x_i\rangle\langle y_i|\}_{i=1}^m$
- *the number i of effects, for which $\{|x_i\rangle, |y_i\rangle\}$ are linearly dependent, is smaller than d.*

Proof. The first condition follows directly from the fact that Kraus operators of $\mathcal{P}_0$ and $\mathcal{P}_1$ are $\{|i\rangle\langle x_i|\}_{i=1}^m$ and $\{|i\rangle\langle y_i|\}_{i=1}^m$ respectively. Therefore, the condition from Theorem 4 can be rewritten as

$$\mathbb{1} \notin \mathrm{span}\{|x_i\rangle\langle i|j\rangle\langle y_j|\}_{i,j=1}^m = \mathrm{span}\{|x_i\rangle\langle y_i|\}_{i=1}^m. \tag{4.1}$$

As for the second condition, assume that the set $\{|x_1\rangle, \ldots, |x_m\rangle\}$ is linearly independent. Similarly, assume that the set $\{|y_1\rangle, \ldots, |y_m\rangle\}$ is linearly independent. Let $|\psi\rangle$ be a quantum state and we define states ρ_0 and ρ_1 as

$$\begin{aligned} \rho_0 &= (\mathcal{P}_0 \otimes \mathbb{1})(|\psi\rangle\langle\psi|) = \sum_i |i\rangle\langle i| \otimes [\psi]^\top |x_i\rangle\langle x_i|^\top \overline{[\psi]}, \\ \rho_1 &= (\mathcal{P}_1 \otimes \mathbb{1})(|\psi\rangle\langle\psi|) = \sum_i |i\rangle\langle i| \otimes [\psi]^\top |y_i\rangle\langle y_i|^\top \overline{[\psi]}. \end{aligned} \tag{4.2}$$

Then $\mathrm{supp}(\rho_0) = \mathrm{span}\{|i\rangle \otimes [\psi]^\top |\tilde{x}_i\rangle\}_i$ and similarly $\mathrm{supp}(\rho_1) = \mathrm{span}\{|i\rangle \otimes [\psi]^\top |\tilde{y}_i\rangle\}_i$, where $[\psi]^\top |\tilde{x}_i\rangle$ is a normalized vector. We want to see whether $\mathrm{supp}(\rho_0) \cap \mathrm{supp}(\rho_1) = \{0\}$. Take any $\rho \in \mathrm{supp}(\rho_0) \cap \mathrm{supp}(\rho_1)$. Then

$$\rho = \sum_i \alpha_i |i\rangle \otimes [\psi]^\top |\tilde{x}_i\rangle = \sum_i \beta_i |i\rangle \otimes [\psi]^\top |\tilde{y}_i\rangle \tag{4.3}$$

from which it follows that

$$\alpha_i [\psi]^\top |\tilde{x}_i\rangle = \beta_i [\psi]^\top |\tilde{y}_i\rangle \tag{4.4}$$

for every i. Therefore, if there does not exist any nonzero c_i such that $|x_i\rangle = c_i |y_i\rangle$, then it must be $\alpha_i = \beta_i = 0$. If there exists c_i such that $|\tilde{x}_i\rangle = c_i |\tilde{y}_i\rangle$ then we need to see whether there exists a state $|\psi\rangle$ such that $[\psi]^\top |\tilde{x}_i\rangle = 0$. A state $|\psi\rangle$ can be orthogonal to up to $d-1$ vectors $|x_i\rangle$. ∎

4.2 Parallel scheme

Parallel scheme is the most straightforward generalization of the single-shot discrimination scheme [35, 102]. In this section, we will focus on the discrimination of quantum measurements, so the schematic representation of the parallel scheme is depicted in Figure 4.1 for the case of quantum measurements.

Let N be the number of queries to the black box, which contains either $\mathcal{P}_0$ or $\mathcal{P}_1$. We prepare an input state $|\psi\rangle$ on $N+1$ registers and perform the measurement contained in the black box on the first N registers. The dimension of the additional $(N+1)$-th register is the same as the dimension of the first N registers together. In other words, we perform the black box measurement N times in parallel. Therefore, we obtain N classical labels and basing on these labels we prepare a final measurement, $\mathcal{P}_F$, on the last register. After this final measurement we make a decision whether the black box contained $\mathcal{P}_0$ or $\mathcal{P}_1$.

Discrimination of measurements $\mathcal{P}_0$ and $\mathcal{P}_1$ in this scenario can be seen as discrimination of tensor products of channels $\mathcal{P}_0 \otimes \ldots \otimes \mathcal{P}_0$ and $\mathcal{P}_1 \otimes \ldots \otimes \mathcal{P}_1$, where

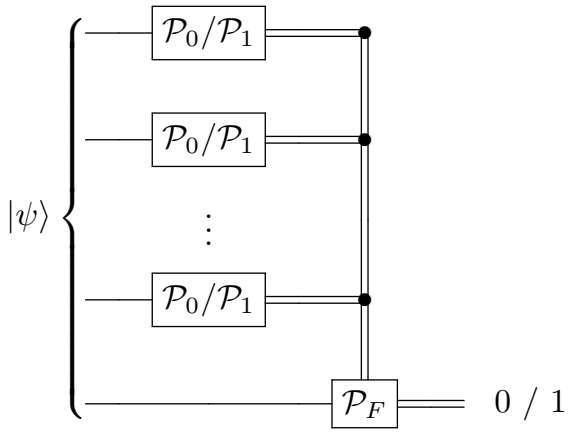

Figure 4.1: Scheme of parallel discrimination of quantum measurements $\mathcal{P}_0$ and $\mathcal{P}_1$.

similarly as in the single-shot case we consider the channels extended by the identity channel. In other words, to bound the probability of correct discrimination in the parallel scheme we can apply the Holevo-Helstrom theorem for tensor products of channels, that is

$$p \leq \frac{1}{2} + \frac{1}{4} \left\| \mathcal{P}_0^{\otimes N} - \mathcal{P}_1^{\otimes N} \right\|_\diamond \tag{4.5}$$

4.3 Adaptive scheme

Adaptive discrimination scheme is a generalization of the parallel scheme. It allows for the use of processing between subsequent queries to the black box. We can adjust the input to the next query to improve the discrimination. The scheme of adaptive discrimination of quantum measurements is presented in Figure 4.2.

For the sake of simplicity we will explain this scheme in the case when the black box is used three times. We prepare an input state $|\psi\rangle$ on $N+1$ registers, which in this case means that our input state is on four registers. Then, on the top register we apply the measurement contained in the black box and as a result we obtain classical label i_1. Next, basing on this label we perform processing Ξ_1 on all the remaining registers. In the next step we perform the measurement hidden in the black box again, this time on the second register. As a result of this measurement we obtain the classical label i_2. Then, we perform the processing Ξ_2 on the third and fourth registers, which this time depends on both labels i_1 and i_2. In what follows, the black box is used again, this time on the third register, and we obtain classical label i_3. Finally, taking into account all the labels i_1, i_2, i_3 we prepare a final measurement, $\mathcal{P}_F$, on the last register. Basing on the outcome of this last measurement we make a decision whether the black box contained either $\mathcal{P}_0$ or $\mathcal{P}_1$.

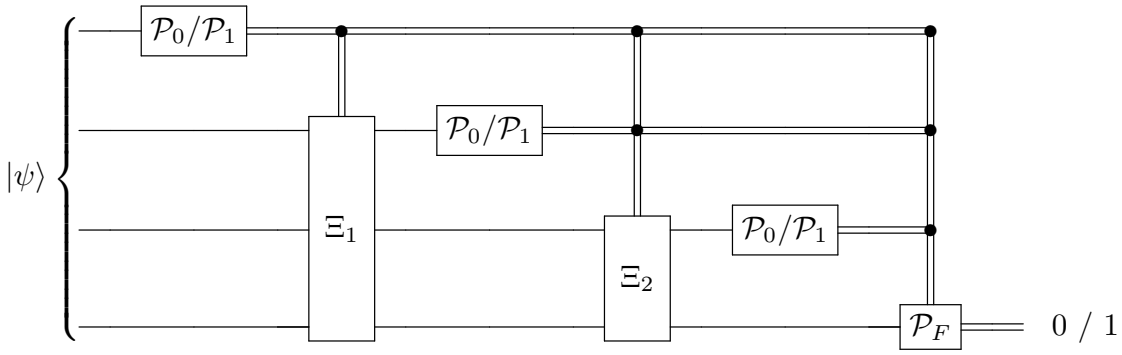

Figure 4.2: Scheme of adaptive discrimination of quantum measurements $\mathcal{P}_0$ and $\mathcal{P}_1$.

The processing in the adaptive scheme can be any quantum channel. We do not make many assumptions on the dimensions of these channels, only we need to make sure that the first register fits the input of the black box in the subsequent query.

A particular example of the processing is the identity channel. In such case, the outcome of the previous measurement in the black box does not impact the input to the next query. This case is equivalent to the parallel discrimination scheme.

4.4 Discrimination of von Neumann measurements

When perfect discrimination is not achievable for a certain pair of von Neumann measurements in the single-shot scenario, it comes as a natural step to use the measurement in the black box many times in the hope that this can improve the discrimination. In this section, we will study both parallel and adaptive discrimination of von Neumann measurements assuming that the black box can be used a finite number of times. We will begin with the parallel scheme and we will calculate the probability of correct discrimination after N queries to the black box in parallel. Later, we will prove that in the case of von Neumann measurements the discrimination cannot be improved by the use of adaptive scheme.

Similarly as in Section 3.3, in this section we will assume that the black box contains one of two von Neumann measurements, either $\mathcal{P}_{\mathbf{1}}$ or $\mathcal{P}_U$, for some given unitary matrix U. In other words, we will be discriminating the measurement in the canonical basis and a measurement in some other basis given by the matrix U.

To calculate the probability of symmetric discrimination between these measurements after N queries in the parallel scheme, we will use the bound from the Holevo-Helstrom theorem in Eq. (4.5). The key part of this bound is the diamond

norm distance between tensor products of the measurements. The following theorem, proved in [35], gives the expression for the diamond norm distance between tensor products of von Neumann measurements.

Theorem 5 *Let $N \in \mathbb{N}$, $U \in \mathcal{U}(\mathcal{X})$ and $\mathcal{P}_U$ be a von Neumann measurement. Then*

$$\|\mathcal{P}_{U^{\otimes N}} - \mathcal{P}_{\mathbf{1}^{\otimes N}}\|_\diamond = \min_{E \in \mathcal{DU}(\mathcal{X})} \|\Phi_{U^{\otimes N} E^{\otimes N}} - \Phi_{\mathbf{1}^{\otimes N}}\|_\diamond . \tag{4.6}$$

Proof. We will consider two cases, when $\|\mathcal{P}_U - \mathcal{P}_{\mathbf{1}}\|_\diamond = 2$, and when $\|\mathcal{P}_U - \mathcal{P}_{\mathbf{1}}\|_\diamond < 2$. We will begin with the former case, when $\|\mathcal{P}_U - \mathcal{P}_{\mathbf{1}}\|_\diamond = 2$, which means that $\mathcal{P}_U$ and $\mathcal{P}_{\mathbf{1}}$ can be discriminated perfectly in a single-shot scheme. Then, from Theorem 2 it holds that

$$\|\mathcal{P}_U - \mathcal{P}_{\mathbf{1}}\|_\diamond = \min_{E \in \mathcal{DU}(\mathcal{X})} \|\Phi_{UE} - \Phi_{\mathbf{1}}\|_\diamond = 2. \tag{4.7}$$

Therefore, for each $N \in \mathbb{N}$ also the measurements $\mathcal{P}_{U^{\otimes N}}$ and $\mathcal{P}_{\mathbf{1}^{\otimes N}}$ can be discriminated perfectly, that is $\|\mathcal{P}_{U^{\otimes N}} - \mathcal{P}_{\mathbf{1}^{\otimes N}}\|_\diamond = 2$. Moreover, it also holds that

$$\begin{aligned} \|\mathcal{P}_{U^{\otimes N}} - \mathcal{P}_{\mathbf{1}^{\otimes N}}\|_\diamond &= \min_{F \in \mathcal{DU}(\mathcal{X}^{\otimes N})} \left\|\Phi_{(U^{\otimes N})F} - \Phi_{\mathbf{1}^{\otimes N}}\right\|_\diamond \\ &\leq \min_{E \in \mathcal{DU}(\mathcal{X})} \|\Phi_{U^{\otimes N} E^{\otimes N}} - \Phi_{\mathbf{1}^{\otimes N}}\|_\diamond \leq 2, \end{aligned} \tag{4.8}$$

which finishes the proof in this case.

Now we will consider the second case when $\|\mathcal{P}_U - \mathcal{P}_{\mathbf{1}}\|_\diamond < 2$, that is when $\mathcal{P}_U$ and $\mathcal{P}_{\mathbf{1}}$ cannot be discriminated perfectly in the single-shot scenario. Then, from Theorem 2 there exists an optimal matrix $E_0 \in \mathcal{DU}(\mathcal{X})$ such that $\|\mathcal{P}_U - \mathcal{P}_{\mathbf{1}}\|_\diamond = \|\Phi_{UE_0} - \Phi_{\mathbf{1}}\|_\diamond$ and $0 \notin W(UE_0)$. In the following part of the proof we will be working towards constructing an input state $\rho_0 \coloneqq \frac{1}{2}\rho_1 + \frac{1}{2}\rho_d$, which will be optimal for discrimination between $\mathcal{P}_{U^{\otimes N}}$ and $\mathcal{P}_{\mathbf{1}^{\otimes N}}$. To do so, we will use Lemma 5 in Appendix A, thus now we will check whether all the assumptions of this Lemma are fulfilled.

Thanks to the equality $\min_{E \in \mathcal{DU}(\mathcal{X})} \|\Phi_{UE} - \Phi_{\mathbf{1}}\|_\diamond = \|\Phi_{UE_0} - \Phi_{\mathbf{1}}\|_\diamond$ we have

$$\max_{E \in \mathcal{DU}(\mathcal{X})} \min_{\rho \in \mathcal{D}(\mathcal{X})} |\mathrm{Tr}(\rho UE)| = \min_{\rho \in \mathcal{D}(\mathcal{X})} |\mathrm{Tr}(\rho UE_0)|. \tag{4.9}$$

Therefore, using the fact that $0 \notin W(UE_0)$, we see that $\min_{\rho \in \mathcal{D}(\mathcal{X})} |\mathrm{Tr}(\rho UE_0)| > 0$. From Lemma 8 in Appendix A, we know that the function $(\rho, E) \mapsto |\mathrm{Tr}(\rho UE)|$ has a saddle point. Therefore, all the assumptions of Lemma 5 are satisfied for the matrix E_0. Let λ_1 and λ_d denote a pair of the most distant eigenvalues of UE_0. From Lemma 5 there exist states $\rho_1, \rho_d \in \mathcal{D}(\mathcal{X})$ such that $\mathrm{diag}(\rho_1) = \mathrm{diag}(\rho_d)$ as

well as $\rho_1 = P_1\rho_1 P_1,\ \rho_d = P_d\rho_d P_d$, where P_1, P_d denote the projectors onto the subspaces spanned by the eigenvectors corresponding to λ_1 and λ_d, respectively.

In the remaining part of the proof we will study two cases: when $\mathcal{P}_U$ and $\mathcal{P}_{\mathbf{1}}$ cannot be discriminated perfectly after N queries, and the case when perfect discrimination is achieved after N steps.

Assume that $\mathcal{P}_U$ and $\mathcal{P}_{\mathbf{1}}$ cannot be discriminated perfectly after N queries, which means that $0 \notin W\left(U^{\otimes N}E_0^{\otimes N}\right)$. Note that as $\mathrm{diag}(\rho_1) = \mathrm{diag}(\rho_d)$, then also

$$\mathrm{diag}(\rho_1^{\otimes N}) = \mathrm{diag}(\rho_d^{\otimes N}). \tag{4.10}$$

Moreover, $\rho_1^{\otimes N}, \rho_d^{\otimes N}$ lie on the subspaces spanned by the eigenvectors corresponding to the eigenvalues λ_1^N and λ_d^N of the matrix $U^{\otimes N}E_0^{\otimes N}$. Therefore, all the latter assumptions of Lemma 5 are fulfilled. From the reverse implication of this Lemma we have that the unitary matrix $E_0^{\otimes N}$ is optimal and for $\rho_0 = \frac{1}{2}\rho_1^{\otimes N} + \frac{1}{2}\rho_d^{\otimes N}$ it holds that

$$\begin{aligned}\min_{\rho\in\mathcal{D}(\mathcal{X})} \left|\mathrm{tr}\left(\rho(UE_0)^{\otimes N}\right)\right| &= \left|\mathrm{tr}\left(\rho_0(UE_0)^{\otimes N}\right)\right| \\ &= \max_{F\in\mathcal{DU}(\mathcal{X}^{\otimes N})} \min_{\rho\in\mathcal{D}(\mathcal{X}^{\otimes N})} \left|\mathrm{tr}\left(\rho U^{\otimes N}F\right)\right|.\end{aligned} \tag{4.11}$$

Hence

$$\begin{aligned}\left\|\Phi_{U^{\otimes N}E_0^{\otimes N}} - \Phi_{\mathbf{1}^{\otimes N}}\right\|_\diamond &= \min_{F\in\mathcal{DU}(\mathcal{X}^{\otimes N})} \left\|\Phi_{U^{\otimes N}F} - \Phi_{\mathbf{1}^{\otimes N}}\right\|_\diamond \\ &\leq \min_{E\in\mathcal{DU}(\mathcal{X})} \left\|\Phi_{U^{\otimes N}E^{\otimes N}} - \Phi_{\mathbf{1}^{\otimes N}}\right\|_\diamond\end{aligned} \tag{4.12}$$

and eventually

$$\left\|\mathcal{P}_{U^{\otimes N}} - \mathcal{P}_{\mathbf{1}^{\otimes N}}\right\|_\diamond = \min_{E\in\mathcal{DU}(\mathcal{X})} \left\|\Phi_{U^{\otimes N}E^{\otimes N}} - \Phi_{\mathbf{1}^{\otimes N}}\right\|_\diamond. \tag{4.13}$$

In the second case, we assume $0 \in W\left(U^{\otimes N}E_0^{\otimes N}\right)$. Let us consider the situation when N is the first index for which this happens, that is $0 \notin W\left(U^{\otimes N-1}E_0^{\otimes N-1}\right)$. Then $0 \in \mathrm{conv}(\lambda_1^N, \lambda_1\lambda_d^{N-1}, \lambda_d^N)$ and there exists a probability vector $p = (p_1, p_2, p_3)$ such that

$$p_1\lambda_1^N + p_2\lambda_1\lambda_d^{N-1} + p_3\lambda_d^N = 0. \tag{4.14}$$

We will show that for the state

$$\rho \coloneqq p_1\rho_1^{\otimes N} + p_2\left(\rho_1 \otimes \rho_d^{\otimes N-1}\right) + p_3\rho_d^{\otimes N}. \tag{4.15}$$

it holds that $\mathrm{diag}\left(\rho U^{\otimes N}\right) = 0$. Indeed

$$\begin{aligned}
&\mathrm{diag}\left(\rho U^{\otimes N} E_0^{\otimes N}\right) \\
&= \mathrm{diag}\left(p_1\lambda_1^N \rho_1^{\otimes N} + p_2\lambda_1\lambda_d^{N-1}\left(\rho_1 \otimes \rho_d^{\otimes N-1}\right) + p_3\lambda_d^N \rho_d^{\otimes N}\right) \\
&= p_1\lambda_1^N \mathrm{diag}(\rho_1^{\otimes N}) + p_2\lambda_1\lambda_d^{N-1}\mathrm{diag}\left(\rho_1 \otimes \rho_d^{\otimes N-1}\right) + p_3\lambda_d^N \mathrm{diag}(\rho_d^{\otimes N}) \\
&= \left(p_1\lambda_1^N + p_2\lambda_1\lambda_d^{N-1} + p_3\lambda_d^N\right)\mathrm{diag}(\rho_1^{\otimes N}) = 0.
\end{aligned} \tag{4.16}$$

Therefore, from Proposition 3 we obtain that $\mathcal{P}_{U^{\otimes N}}$ and $\mathcal{P}_{\mathbf{1}^{\otimes N}}$ can be discriminated perfectly, that is $\|\mathcal{P}_{U^{\otimes N}} - \mathcal{P}_{\mathbf{1}^{\otimes N}}\|_\diamond = 2$ and hence

$$\|\mathcal{P}_{U^{\otimes N}} - \mathcal{P}_{\mathbf{1}^{\otimes N}}\|_\diamond = \min_{E\in\mathcal{DU}(\mathcal{X})} \|\Phi_{U^{\otimes N}E^{\otimes N}} - \Phi_{\mathbf{1}^{\otimes N}}\|_\diamond . \tag{4.17}$$

When M is the first index for which $0 \in W(U^{\otimes M}E_0^{\otimes M})$ and $N > M$, then the equality $\|\mathcal{P}_{U^{\otimes M}} - \mathcal{P}_{\mathbf{1}^{\otimes N}}\|_\diamond = 2$ implies that $\|\mathcal{P}_{U^{\otimes N}} - \mathcal{P}_{\mathbf{1}^{\otimes N}}\|_\diamond = 2$. Therefore,

$$\|\mathcal{P}_{U^{\otimes N}} - \mathcal{P}_{\mathbf{1}^{\otimes N}}\|_\diamond = \min_{E\in\mathcal{DU}(\mathcal{X})} \|\Phi_{U^{\otimes N}E^{\otimes N}} - \Phi_{\mathbf{1}^{\otimes N}}\|_\diamond , \tag{4.18}$$

which completes the proof. ∎

Number of queries for perfect discrimination Let us begin with introducing the notation. Given a unitary matrix $U \in \mathcal{U}(\mathcal{X})$, let $\Theta(U)$ be the angle of the shortest arc containing all eigenvalues of U. We will also take advantage of the optimized version of $\Theta(U)$ defined in Eq. (3.15). Geometrical representation of these quantities was sketched in Figure 3.4 in Sec. 3.3, when we were studying single-shot discrimination of von Neumann measurements. The pair of most distant eigenvalues of the unitary matrix, is denoted on the unit circle by λ_1 and λ_d. The arc between them is signed as thick line and denoted by the symbol Υ. The numerical range of unitary matrix in contained in the gray area. The distance between λ_1 and λ_d corresponds to the diamond norm distance between $\mathcal{P}_U$ and $\mathcal{P}_{\mathbf{1}}$.

Now we are in position to address the question how many times do we need to use the black box to obtain perfect discrimination. Let us recall that if quantum states cannot be discriminated perfectly in the single-shot scheme, they cannot be discriminated perfectly after any finite number of queries [98, 99]. On the other hand, in contrast to discrimination of quantum states, even if quantum channels cannot be discriminated perfectly in the single-shot scenario, in some cases one may achieve perfect discrimination after a finite number of queries. For example, in the case of discrimination of unitary channels Φ_U and $\Phi_{\mathbf{1}}$, perfect discrimination can be always achieved in $N = \lceil \frac{\pi}{\Theta(U)} \rceil$ steps [91].

As for the discrimination of von Neumann measurements in the parallel scheme, let us analyze an example when the black box is used three times. This is sketched

in Figure 4.3.

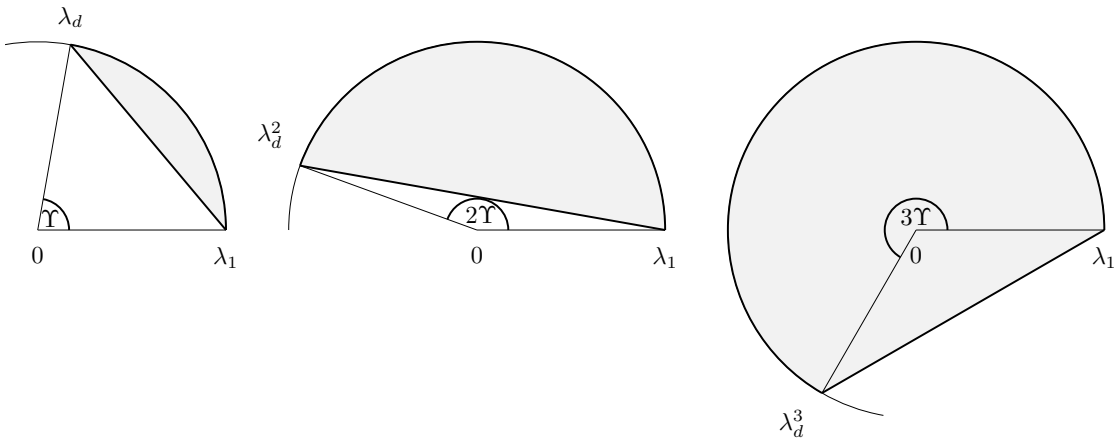

Figure 4.3: Geometrical interpretation of probability of successful discrimination between von Neumann measurements in the parallel scheme. The left, middle and right figures correspond to one, two and three queries, respectively. The numerical ranges of matrices UE_0, $(UE_0)^2$ and $(UE_0)^3$ respectively, are contained in the gray areas.

The left figure corresponds to the first query to the black box. We are discriminating von Neumann measurements $\mathcal{P}_U$ and $\mathcal{P}_{\mathbf{1}}$, and we use the notation as in Fig. 3.4. The most distant eigenvalues of the optimized matrix UE_0 are denoted by $\lambda_1 = 1$ and λ_d. The gray area contains the numerical range of the matrix UE_0. We can see that zero is not included in the numerical range, hence perfect discrimination cannot be achieved after the first query. After the second query we are actually discriminating between $\mathcal{P}_{U^{\otimes 2}}$ and $\mathcal{P}_{\mathbf{1}^{\otimes 2}}$, hence the most distant eigenvalues are now $\lambda_1^2 = \lambda_1 = 1$ and λ_d^2. This is presented in the middle figure. The numerical range is now much bigger, but it still does not contain zero, so perfect discrimination is not obtained. The third query is depicted in the right figure. Now we are discriminating between $\mathcal{P}_{U^{\otimes 3}}$ and $\mathcal{P}_{\mathbf{1}^{\otimes 3}}$, and the most distant eigenvalues are $\lambda_1^3 = 1$ and λ_d^3. This time, finally, the numerical range contains zero and we obtain perfect discrimination between $\mathcal{P}_U$ and $\mathcal{P}_{\mathbf{1}}$ after three queries.

The following proposition gives an expression for the diamond norm distance between tensor products of von Neumann measurements. This allows us to directly calculate the minimal number of queries needed for perfect discrimination between von Neumann measurements.

Proposition 5 *Let $N \in \mathbb{N}$, $U \in \mathcal{U}(\mathcal{X})$. It holds that*

(i) if $N\Upsilon(U) < \pi$, then $\|\mathcal{P}_{U^{\otimes N}} - \mathcal{P}_{\mathbf{1}^{\otimes N}}\|_\diamond = 2\sin\left(\frac{N}{2}\Upsilon(U)\right)$;

(ii) if $N\Upsilon(U) \geq \pi$, *then* $\|\mathcal{P}_{U^{\otimes N}} - \mathcal{P}_{\mathbf{1}^{\otimes N}}\|_\diamond = 2$.

Proof. From Theorem 5 we know that when the matrix UE_0 is in the optimal form, then also the matrix $(UE_0)^{\otimes N}$ is optimal. Hence, to calculating the diamond norm distance between $\mathcal{P}_{U^{\otimes N}}$ and $\mathcal{P}_{\mathbf{1}^{\otimes N}}$ reduces to the problem of determining the value $\Theta((UE_0)^{\otimes N})$.

One can observe that $\Theta((UE_0)^{\otimes N}) = N\Theta(UE_0)$, until $0 \in W\left((UE_0)^{\otimes N}\right)$ (see Fig. 4.3). Therefore, in the case $N\Upsilon(U) < \pi$, the diamond norm distance between $\mathcal{P}_{U^{\otimes N}}$ and $\mathcal{P}_{\mathbf{1}^{\otimes N}}$ equals to the distance between two most distant eigenvalues of $(UE_0)^{\otimes N}$, which is can be expressed as $2\sin\left(\frac{N}{2}\Upsilon(U)\right)$. On the other hand, when $N\Upsilon(U) \geq \pi$, then $0 \in W\left((UE_0)^{\otimes N}\right)$ and the diamond norm distance between $\mathcal{P}_{U^{\otimes N}}$ and $\mathcal{P}_{\mathbf{1}^{\otimes N}}$ equals two.

Moreover, the first time zero enters the numerical range $W\left((UE_0)^{\otimes N}\right)$ is equal to $N = \lceil \frac{\pi}{\Upsilon(U)} \rceil$. ∎

From the above Proposition we can formulate a corollary which provides an expression for the minimal number of queries required to obtain prefect discrimination between von Neumann measurements.

Corollary 5 *The minimal number of queries needed to perfectly discriminate between* $\mathcal{P}_U$ *and* $\mathcal{P}_{\mathbf{1}}$ *in the symmetric setting equals* $N = \lceil \frac{\pi}{\Upsilon(U)} \rceil$.

So far we were studying only the parallel discrimination of von Neumann measurements. We know how to calculate the optimal probability of successful discrimination after N queries and how many queries are needed to achieve perfect discrimination. Let us focus on the case when a pair of measurements can be discriminated perfectly in N queries in parallel. The natural question arises, whether is it possible to achieve perfect discrimination using the black box fewer times, utilizing some processing between subsequent queries. It may seem intuitive that when we can perform processing, we can prepare a better input for the subsequent query to the black box. Hence, we should be able decrease the number of queries to the black box at the cost of additional processing. It turns out that it is not the case, and whenever the measurements require N queries to be discriminated perfectly, the number of queries cannot be decreased by the use of adaptive scheme.

In some cases perfect discrimination cannot be achieved in the parallel scheme. One can ask a question if it is possible to discriminate these measurements perfectly using the adaptive scheme. Or at least, does the adaptive scheme allow for improving the probability of successful discrimination compared to the parallel one. The answer for both questions is negative. The following theorem will prove that the use of adaptive scheme does not give any advantage over the parallel one when discriminating von Neumann measurements.

Theorem 6 *The parallel scheme is optimal for discrimination of von Neumann measurements.*

Proof. Without loss of generality we may assume that the processing is performed using only unitary operations. Indeed, using Stinespring dilation theorem, any channel might be represented via a unitary channel on a larger system followed by the partial trace operation. What is left to observe is that $\|\mathrm{tr}_B(X_{AB})\|_1 \leq \|X_{AB}\|_1$ for arbitrary bipartite matrix X_{AB}.

The sequential scheme of discrimination of von Neumann measurements can be expressed as a channel

$$\Psi_U = (\Delta_{1,\ldots,N} \otimes \mathbb{1})\, \Phi_{A_U}, \tag{4.19}$$

associated with a matrix A_U. Here $\Delta_{1,\ldots,N}$ is the dephasing channel on subsystems $1,\ldots,N$. The channel Φ_{A_U} has the exact form of this transformation given by

$$\begin{aligned} A_U = &(\mathbb{1}_{1,\ldots N-1} \otimes U \otimes \mathbb{1}_{N+1}) \\ &\left(\sum_{i_1,\ldots,i_{N-1}} |i_1,\ldots,i_{N-1}\rangle\langle i_1,\ldots,i_{N-1}| \otimes V^{(N-1)}_{i_1,\ldots,i_{N-1}} \right) \\ &(\mathbb{1}_{1,\ldots,N-2} \otimes U \otimes \mathbb{1}_{N,N+1}) \\ &\left(\sum_{i_1,\ldots,i_{N-2}} |i_1,\ldots,i_{N-2}\rangle\langle i_1,\ldots,i_{N-2}| \otimes V^{(N-2)}_{i_1,\ldots,i_{N-2}} \right) \\ &\ldots \\ &(\mathbb{1}_1 \otimes U \otimes \mathbb{1}_{3,\ldots N+1}) \\ &\left(\sum_{i_1} |i_1\rangle\langle i_1| \otimes V^{(1)}_{i_1} \right) \\ &(U \otimes \mathbb{1}_{2,\ldots N+1})\,. \end{aligned} \tag{4.20}$$

Assuming that matrix U is chosen in the optimal form as in (3.15) *ie.* $\Upsilon(U) = \Theta(U)$ we may calculate the distance between Ψ_U and $\Psi_{\mathbb{1}}$ as

$$\begin{aligned} \max_\rho \|(\Psi_U - \Psi_{\mathbb{1}})(\rho)\|_1 &= \max_\rho \|[(\Delta_{1,\ldots,N} \otimes \mathbb{1})(\Phi_{A_U} - \Phi_{A_{\mathbb{1}}})](\rho)\|_1 \\ &\leq \max_\rho \|(\Phi_{A_U} - \Phi_{A_{\mathbb{1}}})(\rho)\|_1 \\ &\leq \max_\rho \|(\Phi_{U^{\otimes N}\otimes \mathbb{1}} - \Phi_{\mathbb{1}})(\rho)\|_1 \\ &= \|\Phi_{U^{\otimes N}} - \Phi_{\mathbb{1}}\|_\diamond = \|\mathcal{P}_{U^{\otimes N}} - \mathcal{P}_{\mathbb{1}}\|_\diamond\,, \end{aligned} \tag{4.21}$$

where we maximize over states ρ of appropriate dimensions. The induced trace

norm is monotonically decreasing under the action of channels and this gives us the first inequality. The second one follows from the optimality of the parallel scheme of distinguishing unitary channels [100]. Summing up, the adaptive scenario does not give any advantage over the parallel scheme. ■

4.5 Discrimination of SIC POVMs

In the previous section we showed that the adaptive scheme cannot improve the discrimination of von Neumann measurements. But is the parallel discrimination scheme also optimal for more general measurements with rank-one effects? In this section we will see that in some situations the adaptive scheme can significantly improve the discrimination.

We will focus on a class of rank-one measurements having useful symmetry properties, which are SIC POVMs. Discrimination of SIC POVMs in the single-shot case was studied in Section 3.4, hence, in this section we will use the notation introduced in Section 3.4.

We will begin with studying when perfect discrimination can be achieved for a pair of qubit SIC POVMs. It appears that this property can be characterized only by the length of cycles of the permutation. Only permutations which contain a cycle of length 4 can be perfectly discriminated by a finite number of uses in parallel. On top of that, to perfectly discriminate these measurements it suffices to use the black box two times. The optimal states needed for the discrimination will also be stated. Nevertheless, in the case of permutations which do not contain a cycle of length 4, perfect discrimination cannot be achieved after any finite number of queries. The following proposition has not yet been published.

Proposition 6 *Let $\mathcal{P}_0 = \{E_1, \ldots, E_4\}$ and $\mathcal{P}_1 = \{F_1, \ldots, F_4\}$ be SIC POVMs, where $F_i = E_{\pi(i)}$ for every $i = 1, \ldots, 4$. Then*

- *if the permutation π contains a cycle of length 4, then $\mathcal{P}_0$ and $\mathcal{P}_1$ can be discriminated perfectly after two queries in parallel;*
- *if the permutation π does not contain a cycle of length 4, then $\mathcal{P}_0$ and $\mathcal{P}_1$ cannot be discriminated perfectly after any finite number of queries in parallel.*

Proof. We will see when a pair of rank-one measurements can be discriminated perfectly in a finite number of uses by checking conditions from Corollary 4, where $E_i = |x_i\rangle\langle x_i|$ and $F_i = |y_i\rangle\langle y_i|$ for every i. One can directly check that $\mathbb{1} \notin \operatorname{span}\{|x_i\rangle\langle y_i|\}_i$ only if the permutation π contains a cycle of length 4. As for the second condition from Corollary 4, it states that the number of effects for which $|x_i\rangle$ and $|y_i\rangle$ are linearly dependent must be smaller than d. In our case this directly

translates to the number of fixed points of the permutation π. More precisely, this condition is fulfilled if the number of fixed points of the permutation π is smaller than 2. Summing up, if the permutation π contains a cycle of length 4, then $\mathcal{P}_0$ and $\mathcal{P}_1$ can be discriminated perfectly.

It remains to prove that perfect discrimination can be achieved by two uses in parallel. We note that for permutations $(4,3,1,2)$ and $(3,4,2,1)$ the optimal input state for the discrimination is

$$\rho = \frac{1}{3}\begin{bmatrix} 1 & \frac{1}{\sqrt{8}} & \frac{1}{\sqrt{8}} & \frac{1}{2} \\ \frac{1}{\sqrt{8}} & \frac{1}{2} & \frac{1}{2} & -\frac{1}{\sqrt{8}} \\ \frac{1}{\sqrt{8}} & \frac{1}{2} & \frac{1}{2} & -\frac{1}{\sqrt{8}} \\ \frac{1}{2} & -\frac{1}{\sqrt{8}} & -\frac{1}{\sqrt{8}} & 1 \end{bmatrix}.$$

For permutations $(2,4,1,3)$ and $(3,1,4,2)$ the optimal input state is

$$\rho = \frac{1}{3}\begin{bmatrix} 1 & \overline{a} & \overline{a} & \overline{c} \\ a & \frac{1}{2} & \frac{1}{2} & \overline{b} \\ a & \frac{1}{2} & \frac{1}{2} & \overline{b} \\ c & b & b & 1 \end{bmatrix}$$

where $a = \frac{1}{\sqrt{8}}\mathrm{e}^{-\mathrm{i}\frac{2}{3}\pi}$, $b = \frac{1}{\sqrt{8}}\mathrm{e}^{\mathrm{i}\frac{\pi}{3}}$ and $c = \frac{1}{2}\mathrm{e}^{-\mathrm{i}\frac{2}{3}\pi}$, while for permutations $(2,3,4,1)$ and $(4,1,2,3)$ the optimal input state is

$$\rho = \frac{1}{3}\begin{bmatrix} 1 & a & a & c \\ \overline{a} & \frac{1}{2} & \frac{1}{2} & b \\ \overline{a} & \frac{1}{2} & \frac{1}{2} & b \\ \overline{c} & b & b & 1 \end{bmatrix}.$$

■

Now we proceed to studying the parallel discrimination of SIC POVMs for any dimension. Naturally, we assume that SIC POVMs in given dimensions exist. We will focus on the diamond norm distance between a pair of tensor products of SIC POVMs. The following proposition, which is a new result, states the lower bound on this distance.

Proposition 7 *Let $\mathcal{P}_0 = \{E_1, \ldots, E_{d^2}\}$ and $\mathcal{P}_1 = \{F_1, \ldots, F_{d^2}\}$ be SIC POVMs, where $F_i = E_{\pi(i)}$ for every $i = 1, \ldots, d^2$. Let k denote the number of fixed points of the permutation π. Then, the diamond norm distance between $\mathcal{P}_0^{\otimes N}$ and $\mathcal{P}_1^{\otimes N}$ is*

lower-bounded as

$$\left\|\mathcal{P}_0^{\otimes N} - \mathcal{P}_1^{\otimes N}\right\|_\diamond \geq \frac{2}{d^{2N}} \sum_{s=1}^{N} \gamma_{k,N}(s)(d^2-k)^s \sqrt{1 - \frac{1}{(d+1)^s}}, \tag{4.22}$$

where $\gamma_{k,N}(s) = \binom{N}{N-s} k^{N-s}$ *and* $\gamma_{0,N}(N) = 1$. *Moreover, when the permutation* π *does not have fixed points, then*

$$\left\|\mathcal{P}_0^{\otimes N} - \mathcal{P}_1^{\otimes N}\right\|_\diamond \geq 2\sqrt{1 - \frac{1}{(d+1)^N}}. \tag{4.23}$$

Proof. We will calculate the lower bound from Eq (2.52), that is we will calculate

$$\left\|\mathcal{P}_0^{\otimes N} - \mathcal{P}_1^{\otimes N}\right\|_\diamond \geq \frac{1}{d^N} \left\|J\left(\mathcal{P}_0^{\otimes N}\right) - J\left(\mathcal{P}_1^{\otimes N}\right)\right\|_1. \tag{4.24}$$

We begin with stating the Choi matrices of the measurements

$$\begin{aligned} J(\mathcal{P}_0^{\otimes N}) &= \sum_{i_1,\ldots,i_N=1}^{d^2} |i_1 \ldots i_N\rangle\langle i_1 \ldots i_N| \otimes (E_{i_1} \otimes \ldots \otimes E_{i_N})^\top, \\ J(\mathcal{P}_1^{\otimes N}) &= \sum_{i_1,\ldots,i_N=1}^{d^2} |i_1 \ldots i_N\rangle\langle i_1 \ldots i_N| \otimes (F_{i_1} \otimes \ldots \otimes F_{i_N})^\top. \end{aligned} \tag{4.25}$$

Now we calculate

$$\left\|J\left(\mathcal{P}_0^{\otimes N}\right) - J\left(\mathcal{P}_1^{\otimes N}\right)\right\|_1 = \sum_{i_1,\ldots,i_N=1}^{d^2} \|E_{i_1} \otimes \ldots \otimes E_{i_N} - F_{i_1} \otimes \ldots \otimes F_{i_N}\|_1 \tag{4.26}$$

and from Lemma 1 we have

$$\begin{aligned} &\|E_{i_1} \otimes \ldots \otimes E_{i_N} - F_{i_1} \otimes \ldots \otimes F_{i_N}\|_1 \\ &= \left\| \frac{1}{d^N} |\phi_{i_1} \ldots \phi_{i_N}\rangle\langle \phi_{i_1} \ldots \phi_{i_N}| - \frac{1}{d^N} |\phi_{\pi(i_1)} \ldots \phi_{\pi(i_N)}\rangle\langle \phi_{\pi(i_1)} \ldots \phi_{\pi(i_N)}| \right\|_1 \\ &= \sqrt{\left(\frac{2}{d^N}\right)^2 - \frac{4}{d^{2N}} \left|\langle \phi_{i_1} \ldots \phi_{i_N} | \phi_{\pi(i_1)} \ldots \phi_{\pi(i_N)}\rangle\right|^2}. \end{aligned} \tag{4.27}$$

Hence

$$\left\|J(\mathcal{P}_0^{\otimes N}) - J(\mathcal{P}_1^{\otimes N})\right\|_1 = \sum_{i_1,\ldots,i_N=1}^{d^2} \sqrt{\left(\frac{2}{d^N}\right)^2 - \frac{4}{d^{2N}}\left|\langle\phi_{i_1}\ldots\phi_{i_N}|\phi_{\pi(i_1)}\ldots\phi_{\pi(i_N)}\rangle\right|^2}$$
$$= \frac{2}{d^N}\sum_{i_1,\ldots,i_N=1}^{d^2} \sqrt{1 - \left|\langle\phi_{i_1}\ldots\phi_{i_N}|\phi_{\pi(i_1)}\ldots\phi_{\pi(i_N)}\rangle\right|^2} \tag{4.28}$$

The value of $\left|\langle\phi_{i_1}\ldots\phi_{i_N}|\phi_{\pi(i_1)}\ldots\phi_{\pi(i_N)}\rangle\right|^2$ depends on the number of fixed points of the permutation π, that is $\left|\langle\phi_{i_1}\ldots\phi_{i_N}|\phi_{\pi(i_1)}\ldots\phi_{\pi(i_N)}\rangle\right|^2 = \frac{1}{(d+1)^s}$ where s is the number of $\langle\phi_{i_l}|\phi_{\pi(i_l)}\rangle$ for which $|\langle\phi_{i_l}|\phi_{\pi(i_l)}\rangle| \neq 1$.

Finally, the lower bound for diamond norm distance between tensor products of SIC POVMs yields

$$\frac{2}{d^{2N}}\sum_{s=1}^{N}\gamma_{k,N}(s)(d^2-k)^s\sqrt{1-\frac{1}{(d+1)^s}} \tag{4.29}$$

where

$$\gamma_{k,N}(s) = \binom{N}{N-s}k^{N-s} \tag{4.30}$$

and $\gamma_{0,N}(N) = 1$. Moreover, when the permutation π does not have fixed points, then we substitute $k = 0$ and the above simplifies to

$$2\sqrt{1-\frac{1}{(d+1)^N}}. \tag{4.31}$$

■

The following proposition provides conditions when quantum measurements can be discriminated perfectly in the adaptive scheme, but cannot be discriminated perfectly after any finite number of steps in the parallel scheme.

Proposition 8 *Let $\mathcal{P}_0$ and $\mathcal{P}_1$ be quantum measurements of dimension d with effects $\{|x_i\rangle\langle x_i|\}_i$ and $\{|y_i\rangle\langle y_i|\}_i$ respectively. If*

- $\mathbb{1} \notin \mathrm{span}\{|x_i\rangle\langle y_i|\}_i$,
- *the number of effects for which $|x_i\rangle$ and $|y_i\rangle$ are linearly dependent is smaller than d.*
- $\mathrm{span}\{|x_i\rangle\langle y_i|\}_i$ *contains a positive operator $\rho > 0$,*

then $\mathcal{P}_0$ and $\mathcal{P}_1$ can be discriminated perfectly only by the adaptive scheme.

The first and second conditions in the above Proposition come from the Corollary 4. They assure that quantum measurements can be discriminated perfectly after a finite number of queries to the black box. The last condition corresponds to the fact that the parallel scheme is not sufficient for perfect discrimination and one needs to utilize the adaptive scheme [102].

The first example of quantum channels, which require the adaptive strategy to be perfectly discriminated, was introduced in [103]. Now we can see that many pairs of quantum measurements fulfill the conditions of the above Corollary and require adaptive strategies to be discriminated perfectly. We will construct an exemplary pair of such measurements and present a detailed scheme of adaptive discrimination [34]. It should come as no surprise that this example will be related with SIC POVMs. However, they cannot be of dimension two. Recall from Proposition 6 that a pair of qubit SIC POVMs can be either discriminated perfectly after two queries in parallel or they cannot be discriminated perfectly after any finite number of queries. Hence, to find a good example, we will look into SIC POVMs of dimension three.

Measurements which can be discriminated only adaptively. Now we will construct an example of measurements which can be discriminated perfectly only in the adaptive scheme. We will use the construction of qutrit SIC POVMs from [104]. Given a matrix

$$\frac{1}{\sqrt{2}}\begin{bmatrix} 0 & 0 & 0 & -1 & -\omega_3 & -\omega_3^2 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 & -1 & -\omega_3 & -\omega_3^2 \\ -1 & -\omega_3 & -\omega_3^2 & 1 & 1 & 1 & 0 & 0 & 0 \end{bmatrix} \tag{4.32}$$

where $\omega_3 = \exp\left(\frac{2\pi \mathrm{i}}{3}\right)$, we define the SIC POVM $\mathcal{P}_0$ as a quantum measurement having effects $\{|x_i\rangle\langle x_i|\}_i$, where $|x_i\rangle$ is the i-th column of the matrix (4.32). Let $\pi = (9, 8, 7, 3, 1, 2, 6, 4, 5)$ be a permutation. We define $\mathcal{P}_1$ as a SIC POVM with effects $\{|y_i\rangle\langle y_i|\}_i$, where $|y_i\rangle = |x_{\pi(i)}\rangle$ for $i = 1, \ldots, 9$.

Let us verify that $\mathcal{P}_0$ and $\mathcal{P}_1$ indeed require adaptive scheme to be discriminated perfectly. We will check the conditions from Proposition 8. The first condition can be checked directly. The second condition is trivially fulfilled due to the SIC condition and the fact that the permutation π has no fixed points. As for the third condition, that is the existence of positive operator in $\mathrm{span}\{|x_i\rangle\langle y_i|\}_i$, it can be checked by the following iterative algorithm which was introduced in Section 5 in [34].

Algorithm for finding a positive operator (last condition in Proposition 8)

(i) Construct a projection operator P on the space $\operatorname{span}\{|x_i\rangle\langle y_i|\}_{i=1}^m$ and choose an initial operator $X \in \mathcal{L}(\mathcal{X})$.

(ii) Project X onto the subspace given by P, obtaining some operator Y.

(iii) Substitute the operator X with the quantum state closest to Y.

(iv) Repeat the procedure from points (ii) and (iii) until it converges or a predefined number of steps is achieved.

(v) Verify if the obtained operator is of full rank.

Scheme of adaptive discrimination Now we will focus on describing the scheme of adaptive discrimination between SIC POVMs $\mathcal{P}_0$ and $\mathcal{P}_1$ in greater detail. The adaptive procedure is depicted in Figure 4.4 and for the sake of clarity, we will be denoting the space $\mathcal{X}$ as corresponding to the top register and the spaces $\mathcal{Y}$, $\mathcal{Z}$ will correspond to second and third registers, respectively. As both $\mathcal{P}_0$ and $\mathcal{P}_1$ are qutrit SIC POVMs, all the spaces $\mathcal{X}, \mathcal{Y}, \mathcal{Z}$ are actually equal $\mathbb{C}^3$.

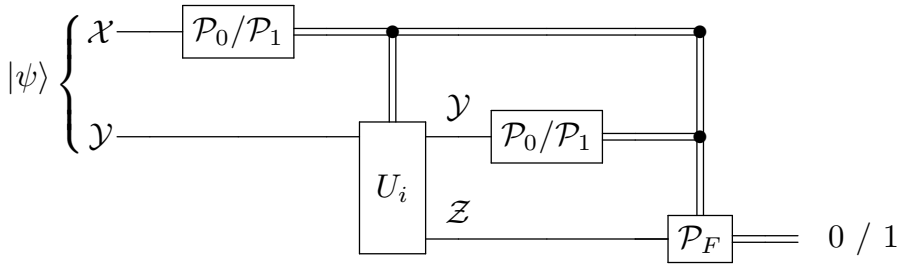

Figure 4.4: Scheme of adaptive discrimination between SIC POVMs $\mathcal{P}_0$ and $\mathcal{P}_1$.

Before presenting the algorithm, let us introduce some additional notation. Let $|\psi_i\rangle \coloneqq \frac{|x_i\rangle}{\||x_i\rangle\|}$ and $|\varphi_i\rangle \coloneqq \frac{|y_i\rangle}{\||y_i\rangle\|}$ for every $i = 1, \ldots, 9$.

Let $A \in \mathrm{L}(\mathcal{X})$ be a nonzero matrix satisfying

$$A \perp \operatorname{span}\{|x_i\rangle\langle y_i|\}_i. \tag{4.33}$$

Note that such a matrix exists due to the fact that $\operatorname{rank}(\operatorname{span}\{|x_i\rangle\langle y_i|\}_i) < 9$. Let $A = U\Sigma V^\dagger$ be the singular value decomposition of this matrix. (See Eq. (2.8), where Σ is a diagonal matrix of singular values and matrices U and V form orthonormal bases.)

We define vectors $|\tilde{\xi}\rangle, |\tilde{\eta}\rangle \in \mathcal{Y} \otimes \mathcal{Z}$ as

$$|\tilde{\xi}\rangle \coloneqq |U\sqrt{\Sigma}\rangle\!\rangle, \quad |\tilde{\eta}\rangle \coloneqq |V\sqrt{\Sigma}\rangle\!\rangle \tag{4.34}$$

and see that they fulfill the condition

$$\mathrm{Tr}_{\mathcal{Z}}|\tilde{\xi}\rangle\langle\tilde{\eta}| = A.$$

We also define the normalized vectors $|\xi\rangle \coloneqq \frac{|\tilde{\xi}\rangle}{\||\tilde{\xi}\rangle\|}$ and $|\eta\rangle \coloneqq \frac{|\tilde{\eta}\rangle}{\||\tilde{\eta}\rangle\|}$.

Finally, for every $i = 1, \ldots, 9$ we define an isometric channel $\Phi_{U_i}(\cdot)$ for $U_i \in \mathrm{U}(\mathcal{Y}, \mathcal{Y} \otimes \mathcal{Z})$ given by the conditions

$$\begin{aligned} U_i|\psi_i\rangle &= \mathrm{e}^{\mathrm{i}\theta_i}|\xi\rangle \\ U_i|\varphi_i\rangle &= |\eta\rangle \end{aligned} \tag{4.35}$$

where $\theta_i = \beta - \alpha_i$ for $\langle\psi_i|\varphi_i\rangle = r_i\mathrm{e}^{\mathrm{i}\alpha_i}$ and $\langle\xi|\eta\rangle = r\mathrm{e}^{\mathrm{i}\beta}$.

Now, we are in the position to present the algorithm for the adaptive discrimination between $\mathcal{P}_0$ and $\mathcal{P}_1$.

Algorithm for adaptive discrimination between SIC POVMs.

1. Prepare an input state $|\psi\rangle = \frac{1}{\sqrt{d}}|\mathbb{1}_{\mathcal{X}}\rangle\!\rangle$ on the first and second registers.
2. Perform the unknown measurement, either $\mathcal{P}_0$ or $\mathcal{P}_1$, on the first register, and obtain a label i. Note that at the same time on the second register the state is either $|\psi_i\rangle$ or $|\varphi_i\rangle \in \mathcal{Y}$.
3. Based on the label i, perform an isometric channel Φ_{U_i} on the second register. After performing this isometric channel there is a new – third register $\mathcal{Z}$.
4. Perform the unknown measurement on the (second) register $\mathcal{Y}$ and obtain a label j.
5. Finally, measure the third register and make a decision whether in the black box there was either $\mathcal{P}_0$ or $\mathcal{P}_1$.

So far we know how to adaptively discriminate pairs of SIC POVMs of dimension three. One may ask a question if the same thing can be done more generally, for higher dimensions. Although the algorithm is described for dimension three, nothing stands in the way to apply it in higher dimensions. One can numerically check that in dimension four there also exist pairs of SIC POVMs which can be perfectly discriminated only by the use of adaptive scheme. For such pairs one can also use the above algorithm.

4.6 Discrimination of general rank-one POVMs

We already know that in the case of von Neumann measurements the parallel scheme is optimal for discrimination. However, in the previous section we saw an example where the use of the adaptive scheme significantly improves the discrimination. More precisely, the adaptive scheme allowed for perfect discrimination of measurements which could not be discriminated perfectly after any finite number of queries in the parallel scheme.

From the above paragraph we can see that when rank-one measurements of dimension d have d effects, then the parallel scheme is always sufficient for discrimination. That was the case of von Neumann measurements. On the other hand, when the rank-one measurements have d^2 effects, then the adaptive scheme may improve the discrimination. So what happens if the rank-one measurements have $d < m < d^2$ effects? We will study this problem numerically. We will generate random rank-one measurements with m effects and check whether pairs of the generated measurements require adaptive strategies for perfect discrimination.

Generating random rank-one POVMs. Let describe the procedure of generating random measurements of dimension d with $d < m < d^2$ rank-one effects $\{|x_i\rangle\langle x_i|\}_{i=1}^m$. We begin with generating a Haar-random isometry of dimensions $d \times m$. [105,106] Then, we take projectors onto the columns $\{|x_i\rangle\}_{i=1}^m$ of this matrix. The obtained projections $\{|x_i\rangle\langle x_i|\}_{i=1}^m$ are the effects of the generated rank-one measurement.

As we know how to generate random rank-one measurements, we need be able to check when adaptive scheme improves the discrimination. We will focus on the case when this pair of measurements cannot be discriminated perfectly using the parallel scheme, but the adaptive scheme allows for perfect discrimination. To see when this is the case, we need to check if the conditions from the Proposition 8 are fulfilled.

Assume we have taken at random a pair of measurements, say $\mathcal{P}_0$ and $\mathcal{P}_1$, with effects $\{|x_i\rangle\langle x_i|\}_{i=1}^m$ and $\{|y_i\rangle\langle y_i|\}_{i=1}^m$, respectively. We will go through each point of the Proposition 8 and see how to check them for our random pair of rank-one measurements.

The first condition requires verifying if the space $\mathrm{span}\{|x_i\rangle\langle y_i|\}_i$ does not contain the identity operator. As our measurements are generated according to the above algorithm, we know that any pair of $|x_i\rangle$ and $|x_j\rangle$ is orthogonal. Hence, if $m = d^2$, then $\mathbb{1} \in \mathrm{span}\{|x_i\rangle\langle y_i|\}_i$, while for $d \leq m < d^2$ it holds that $\mathbb{1} \notin \mathrm{span}\{|x_i\rangle\langle y_i|\}$. The second condition requires that the number of effects for which $|x_i\rangle$ and $|y_i\rangle$ are linearly independent is smaller than d. For the measurements generated in the way described above, this condition is always fulfilled. In the third condition we need

to check if $\text{span}\{|x_i\rangle\langle y_i|\}_{i=1}^m$ contains a positive operator. To to this, we can utilize the iterative algorithm introduces in the previous section after Proposition 8.

Summing up, the first condition is fulfilled as long as $m < d^2$. The second condition is always fulfilled for rank-one measurements. Therefore, to verify if the adaptive scheme improves discrimination between the generated pairs of measurements, we need to numerically check the last condition, that is use the algorithm for finding the positive operator in $\text{span}\{|x_i\rangle\langle y_i|\}_{i=1}^m$.

Simple numerical calculations show that for dimension two there exist pairs of rank-one measurements which fulfill the third condition from Proposition 8. These measurements have three effects. Additionally, it turns out that in the qubit case the probability of finding a pair of POVMs which require adaptive scheme to be perfectly discriminated is roughly equal to $\frac{4}{10}$.

As we know the qubit case, what about higher dimensions? Does adaptive scheme improves the discrimination? How does it depend on the number of effects? We studied this problem numerically for dimension $d = 7$. We sampled $N = 10^6$ pairs of random rank-one POVMs according to the algorithm described at the beginning of this section. These measurements have $d < m < d^2$ effects. For each pair of measurements we checked the third condition from Proposition 8 using the algorithm for finding the positive operator in $\text{span}\{|x_i\rangle\langle y_i|\}_{i=1}^m$ which was introduces in the previous section. Figure 4.5 shows the numerically calculated probability that this condition is fulfilled, as a function of the number of number of measurement's effects [34].

We can see in the figure that as the number of effects increases, so does the probability that the adaptive scheme is needed for perfect discrimination. This agrees with the results discussed earlier in this chapter. When $m = d$, then we have von Neumann measurements and they can always be discriminated perfectly in the parallel scheme. On the other hand, when $m = d^2$, then $\text{span}\{|x_i\rangle\langle y_i|\}_i$ gives the entire space $\mathbb{C}^d$ and we are guaranteed to find the identity operator in it. In such case, random measurements with d^2 effects cannot be discriminated perfectly after any finite number of uses. Note that the situation is a bit different in the case of SIC POVMs as they fulfill the SIC condition in Eq. (3.16). In general, it is clear that we are more likely to need the adaptive scheme for perfect discrimination when the number of effects is closer to d^2.

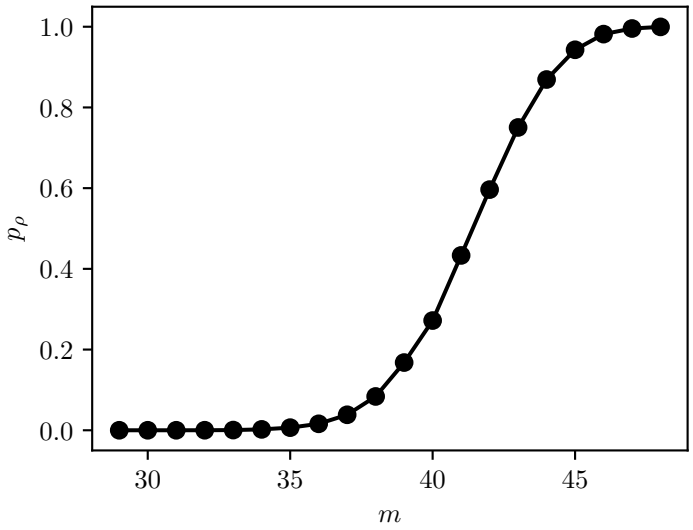

Figure 4.5: Numerically estimated lower bound for the probability p_ρ of an event that for two randomly sampled POVMs with rank-one effects the adaptive scenario is necessary for achieving perfect discrimination. The probability is plotted as a function of the number of effects m for the dimension $d = 7$. The number of samples per point is $N = 10^6$.

Chapter 5

Unambiguous discrimination

In many real-life situations making mistakes can be very costly. When dealing with valuable objects, we do not want to confound them. A similar situation concerns the discrimination of quantum objects. In some situations, we may be specifically interested in discriminating them and being sure that we do not make a mistake. The situation is simple when quantum channels or measurements can be discriminated perfectly. Unfortunately, in most situations this is not the case. Then we have two options. We can either perform the protocol of symmetric discrimination, thus minimize the probability of making an erroneous decision. The major drawback of this approach is that whenever we get the result of discrimination, this result is correct only up to some probability. The second option is to use another approach to the discrimination problem known as the *unambiguous* discrimination. Unambiguous discrimination was first studied for discrimination of quantum states [107–112], and later also for discrimination of quantum channels [82, 90, 113] and measurements [35, 114].

The unambiguous discrimination of two objects allows for three possible outcomes of the discrimination. Let us consider the standard situation when we are discriminating either measurements or channels. One of them is secretly chosen and hidden in the black box with equal probabilities. After performing the protocol of discrimination, we can either know which object was in the black box or we may get an inconclusive answer. When we know which object was hidden in the black box, we know it with certainty. Therefore, we cannot make a mistake and mix the two objects. However, there is no such thing as a free lunch. There is a chance that the discrimination protocol will not indicate which object was in the black box; that is, we will obtain an inconclusive result. Nevertheless, when the chance of getting the inconclusive result is sufficiently small, it may be worth taking this risk, as thanks to that, we can be sure that we will not mix the quantum objects in the black box with one another.

This chapter will focus on the unambiguous discrimination of quantum mea-

surements and will be based on [35]. We will begin with introducing the general entanglement-assisted scheme of unambiguous discrimination of quantum measurements in Section 5.1. We will formulate how to calculate the probability of successful discrimination and explain what it actually means in the unambiguous setup.

Later, we will study the single-shot discrimination of measurements with rank-one effects in Section 5.2. We will prove an upper bound on the probability of successful unambiguous discrimination of general rank-one measurements. We will also state similar bounds for the discrimination of von Neumann measurements and SIC POVMs. On top of that, we will present a geometrical interpretation for the probability of unambiguous discrimination of von Neumann measurements.

The following Section 5.3 will concern the special case of single-shot discrimination, that is, the discrimination without entanglement. We will assume that no additional register can be used and calculate an upper bound on the probability of successful unambiguous discrimination.

Is unambiguous discrimination of quantum measurements always possible? Or maybe in some cases it does not make any sense to even try this scheme? We will try to answer these questions in Section 5.4. We will state the conditions for unambiguous discrimination of general quantum channels and measurements.

Finally, in Section 5.5, we will explore the problem of multiple-shot unambiguous discrimination. We will introduce the parallel and adaptive discrimination schemes and state the probability of unambiguous discrimination after N queries to the black box in the parallel scheme. We will also see the geometrical interpretation of this quantity for the case of von Neumann measurements. Eventually, we will study whether the adaptive scheme can improve the discrimination.

5.1 Scheme of unambiguous discrimination

In this section, we will introduce the general scheme of unambiguous discrimination of quantum measurements. We will focus on discrimination between two quantum measurements $\mathcal{P}_0$ and $\mathcal{P}_1$. Assume that one of these measurements is hidden in the black box and we want to discriminate them unambiguously using the discrimination scheme depicted in Figure 5.1.

We prepare an input state $|\psi\rangle$ on both registers and perform the measurement hidden in the black box on the first register. We know that in the black box there was either the measurement $\mathcal{P}_0$, or $\mathcal{P}_1$. After applying one of these measurements, we obtain a classical label on the first register. Basing on this label we prepare a final measurement and apply it on the second register. This final measurement, $\mathcal{P}_F$, has three effects, which are $\{\Omega_0, \Omega_1, \Omega_?\}$. When we obtain the label associated with the effect Ω_0, then we are sure that the channel in the black box was $\mathcal{P}_0$.

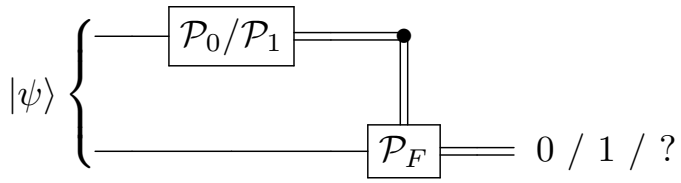

Figure 5.1: Scheme of unambiguous discrimination of quantum measurements $\mathcal{P}_0$ and $\mathcal{P}_1$.

Analogously, if we obtain the label associated with the effect Ω_1, then we are sure that the channel in the black box was $\mathcal{P}_1$. Nevertheless, there is a third possibility that the label will be associated with the effect $\Omega_?$. In such a case we do not know which of the channels was hidden in the black box. This third option will be called an *inconclusive* answer.

What do we mean when saying that unambiguous discrimination was successful? In the above scheme, the probability of successful unambiguous discrimination between quantum measurements $\mathcal{P}_0$ and $\mathcal{P}_1$ equals

$$p_u(\mathcal{P}_F, \psi) = \frac{1}{2}\mathrm{tr}\left(\Omega_0 \left(\mathcal{P}_0 \otimes \mathbb{1}\right)(\psi)\right) + \frac{1}{2}\mathrm{tr}\left(\Omega_1 \left(\mathcal{P}_1 \otimes \mathbb{1}\right)(\psi)\right). \tag{5.1}$$

This probability depends on the choice of the input state ψ and final measurement $\mathcal{P}_F$. The first part of the right hand side of the above formula corresponds to the situation when in the black box there was $\mathcal{P}_0$ (with a priori probability equal $1/2$), and the outcome of the final measurement was associated with the effect Ω_0. Similarly, the second part correspond to the case when the black box contained the measurement $\mathcal{P}_1$ and the label of the final measurement was associated with the effect Ω_1.

The unambiguity condition states that whenever we make a decision which of the measurements was in the black box, this decision is correct. Formally, this can be written as

$$\mathrm{tr}\left(\Omega_0 \left(\mathcal{P}_1 \otimes \mathbb{1}\right)(\psi)\right) = \mathrm{tr}\left(\Omega_1 \left(\mathcal{P}_0 \otimes \mathbb{1}\right)(\psi)\right) = 0. \tag{5.2}$$

Note that third effect, $\Omega_?$, appears neither in the probability of unambiguous discrimination in Eq. (5.1) nor in the unambiguity condition in Eq. (5.2). Intuitively, we say that the unambiguous discrimination was successful when we got the conclusive result. In other words, in unambiguous discrimination we always need to assume the unambiguity condition in Eq. (5.2) and we are interested in the probability that the final measurement gives the label which is not associated with the effect $\Omega_?$.

The *strategy* of unambiguous discrimination involves the choices of the input

state and the final measurement. Recall the probability of unambiguous discrimination in Eq. (5.1). This probability does depend on the chosen input and final measurement. However, we will often be interested in the optimized version of such probability, which does not depend on the discrimination strategy. Thus we introduce

$$p_u \coloneqq \max_{\mathcal{P}_F,\psi} p_u(\mathcal{P}_F,\psi). \tag{5.3}$$

We will sometimes write $p_u(\mathcal{P}_0,\mathcal{P}_1)$ instead of p_u to emphasize which measurements are being discriminated.

5.2 Single-shot discrimination

In this section we will study the optimal probability of unambiguous discrimination in the scheme described in Section 5.1. We will begin with stating a general theorem concerning the discrimination of POVMs with rank-one effects. The presented version of the theorem will be a generalization of the theorem proved in [35] for unambiguous discrimination of von Neumann measurements. Later, we will present two corollaries which give formulas for this probability in the cases of von Neumann measurements and SIC POVMs.

Theorem 7 *The optimal success probability of unambiguous discrimination between rank-one measurements $\mathcal{P}_0$ and $\mathcal{P}_1$ with effects $\{|x_i\rangle\langle x_i|\}_{i=1}^m$ and $\{|y_i\rangle\langle y_i|\}_{i=1}^m$, respectively, is upper-bounded by*

$$p_{\mathrm{u}}(\mathcal{P}_0,\mathcal{P}_1) \leq 1 - \min_{\rho\in\mathcal{D}(\mathcal{X})} \sum_i |\langle x_i|\rho|y_i\rangle|. \tag{5.4}$$

Proof. Let us consider an input state $|\psi\rangle$ on the compound register. Let X be a matrix satisfying $|\psi\rangle = \sum_{i,j} X_{i,j}|i\rangle|j\rangle$.

We perform one of the channels, either $\mathcal{P}_0$ or $\mathcal{P}_1$, extended by the identity channel on the state $|\psi\rangle\langle\psi|$. As a result we obtain one of the states

$$\begin{aligned}(\mathcal{P}_0\otimes\mathbb{1})(|\psi\rangle\langle\psi|) &= \sum_{i=1}^m |i\rangle\langle i|\otimes X^T|x_i\rangle\langle x_i|\overline{X},\\ (\mathcal{P}_1\otimes\mathbb{1})(|\psi\rangle\langle\psi|) &= \sum_{i=1}^m |i\rangle\langle i|\otimes X^T|y_i\rangle\langle y_i|\overline{X}.\end{aligned} \tag{5.5}$$

As both output states have block-diagonal structure, we will restrict our atten-

tion to considering measurements having block-diagonal structure of the form

$$\Omega := \sum_{i=1}^{m} |i\rangle\langle i| \otimes \Omega_i. \tag{5.6}$$

If the obtained label was i, then the state of the auxiliary subsystem is either

$$|a_i\rangle\langle a_i| = a_i^{-1} X^\top (|x_i\rangle\langle x_i|)^\top \overline{X}, \tag{5.7}$$

if the measurement $\mathcal{P}_0$ was performed, or it is

$$|b_i\rangle\langle b_i| = b_i^{-1} X^\top (|y_i\rangle\langle y_i|)^\top \overline{X} \tag{5.8}$$

if the measurement $\mathcal{P}_1$ was performed. The scalars a_i, b_i are responsible for normalization. If any of these scalars was equal zero, then the corresponding outcome would not occur. Therefore, we will assume that $a_i > 0$ and $b_i > 0$. Utilizing the fact that the obtained states $|a_i\rangle\langle a_i|, |b_i\rangle\langle b_i|$ are pure, we can define the final measurements as $\mathcal{P}_F^{(i)} := \{\Omega_0^{(i)}, \Omega_1^{(i)}, M_?^{(i)}\}$, where

$$\begin{aligned} \Omega_0^{(i)} &:= \gamma_0^{(i)}(\mathbb{1} - |b_i\rangle\langle b_i|), \\ \Omega_1^{(i)} &:= \gamma_1^{(i)}(\mathbb{1} - |a_i\rangle\langle a_i|), \\ \Omega_?^{(i)} &:= \mathbb{1} - \Omega_0^{(i)} - \Omega_1^{(i)}, \end{aligned} \tag{5.9}$$

and $\gamma_0^{(i)}, \gamma_1^{(i)}$ guarantee the non-negativity of $\Omega_?^{(i)}$.

The success probability in unambiguous discrimination of pure states $|a\rangle, |b\rangle$ with a priori probabilities $\eta, 1-\eta$ is given by [115]

$$p_{succ}^u(a, b, \eta) = \begin{cases} 1 - \eta - (1-\eta)c^2 & \text{for } \eta < \frac{c^2}{1+c^2} \\ 1 - 2c\sqrt{\eta(1-\eta)} & \text{for } \frac{c^2}{1+c^2} \leq \eta \leq \frac{1}{1+c^2} \\ 1 - (1-\eta) - \eta c^2 & \text{for } \frac{1}{1+c^2} < \eta, \end{cases} \tag{5.10}$$

where $c = |\langle a|b\rangle|$. We will focus on the following upper bound

$$p_{succ}^u(a, b, \eta) \leq 1 - 2c\sqrt{\eta(1-\eta)}. \tag{5.11}$$

Let us define $\rho = XX^\dagger$. When the label i is obtained, the a priori probabilities of obtaining the states $|a_i\rangle, |b_i\rangle$ are $\eta_i = \frac{a_i}{a_i + b_i}, 1 - \eta_i = \frac{b_i}{a_i + b_i}$. The overlap between $|a_i\rangle$ and $|b_i\rangle$ can be calculated as

$$c_i = |\langle a_i|b_i\rangle| = \frac{|\langle \overline{x_i}|\overline{X}X^\top|\overline{y_i}\rangle|}{\sqrt{a_i b_i}} = \frac{|\langle x_i|\rho|y_i\rangle|}{\sqrt{a_i b_i}}. \tag{5.12}$$

Now we are in position to calculate the upper bound for the probability of successful unambiguous discrimination when the label i was observed.

$$p^u_{succ}(a_i, b_i, \eta_i) \leq 1 - 2c_i \frac{\sqrt{a_i b_i}}{a_i + b_i} = 1 - \frac{2|\langle x_i|\rho|y_i\rangle|}{a_i + b_i}. \tag{5.13}$$

Finally, we calculate the bound of the overall success probability as

$$\begin{aligned} p_{\mathrm{u}}(\mathcal{P}_0, \mathcal{P}_1) &= \max_{|\psi\rangle} \sum_i \Pr(\text{label} = i) p^u_{succ}(a_i, b_i, \eta_i) \\ &\leq \max_{\rho} \sum_i \frac{1}{2}(a_i + b_i) \left(1 - \frac{2|\langle x_i|\rho|y_i\rangle|}{a_i + b_i}\right) \\ &= \max_{\rho} \left(\sum_i \frac{1}{2}(a_i + b_i) - \sum_i |\langle x_i|\rho|y_i\rangle|\right) \\ &= 1 - \min_{\rho} \sum_i |\langle x_i|\rho|y_i\rangle|. \end{aligned} \tag{5.14}$$

■

Let us now consider two classes of measurements with rank-one effects, which were also of special importance in the previous chapters. Theses classes are von Neumann measurements and SIC POVMs. As for the von Neumann measurements, similarly as in the symmetric discrimination, we can assume that one of the measurements is in the canonical basis. Therefore, we can restrict our attention to the problem of discrimination between $\mathcal{P}_{\mathbf{1}}$ and $\mathcal{P}_U$, for some unitary matrix U. To apply the expression for unambiguous discrimination stated in Theorem 7 for the case of von Neumann measurements, we need to known their Kraus operators. Fortunately, we can directly see that the Kraus operators of $\mathcal{P}_{\mathbf{1}}$ are $\{|i\rangle\langle i|\}_i$ while the Kraus operators of $\mathcal{P}_U$ are $\{U|i\rangle\langle i|U^\dagger\}_i$. Hence, we have the following corollary.

Corollary 6 *The success probability of unambiguous discrimination between von Neumann measurements $\mathcal{P}_{\mathbf{1}}$ and $\mathcal{P}_U$ is upper-bounded by*

$$p_{\mathrm{u}}(\mathcal{P}_{\mathbf{1}}, \mathcal{P}_U) \leq 1 - \min_{\rho \in \mathcal{D}(\mathcal{X})} \sum_i |\langle i|\rho U|i\rangle|. \tag{5.15}$$

The bound in the above Corollary is tight. When $\mathcal{P}_{\mathbf{1}}$ and $\mathcal{P}_U$ are can be discriminated perfectly, we can utilize Proposition 12 in Appendix A. In the case when $\mathcal{P}_{\mathbf{1}}$ and $\mathcal{P}_U$ cannot be discriminated perfectly, then it follows from Lemma 5 in Appendix A.

Figure 5.2 represents a geometrical interpretation of the probability of unambiguous discrimination between von Neumann measurements from the above

Corollary. The notation used is the same as in Figure 3.4 in Section 3.3. The value of the probability of unambiguous discrimination is denoted by p_u, while λ_1 and λ_d denote a pair of the most distant eigenvalues of the optimized unitary matrix UE_0 (see Eq. (3.15)).

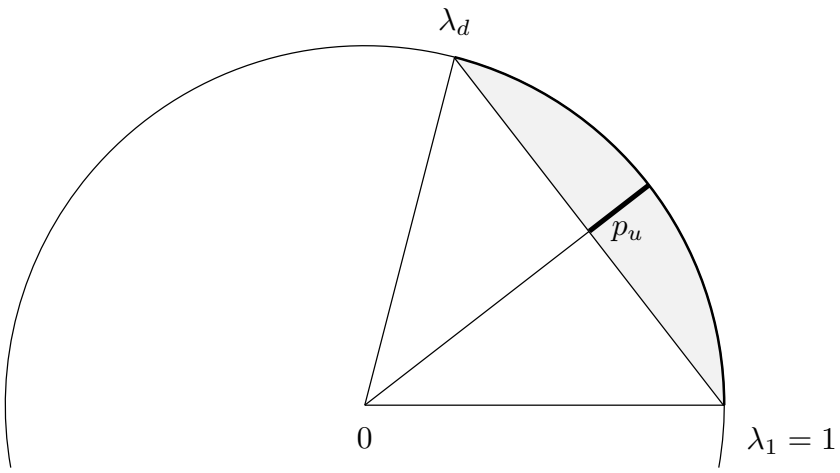

Figure 5.2: Geometrical interpretation of the probability p_u of unambiguous discrimination between von Neumann measurements $\mathcal{P}_U$ and $\mathcal{P}_{\mathbb{1}}$. λ_1 and λ_d are the most distant eigenvalues of the unitary matrix UE_0 and the numerical range of this matrix is contained in the gray area.

Now we proceed to unambiguous discrimination of SIC POVMs. Let $\mathcal{P}_0$ be a SIC POVM with effects $\{|x_i\rangle\langle x_i|\}_{i=1}^{d^2}$ and let $\mathcal{P}_1$ has the effects $\{|y_i\rangle\langle y_i|\}_{i=1}^{d^2}$, where $|y_i\rangle = |x_{\pi(i)}\rangle$ and π is a permutation of d^2 numbers. The following Corollary states the upper bound for the unambiguous discrimination depending of on the number of fixed points of the permutation π.

Corollary 7 *The success probability of unambiguous discrimination between SIC POVMs $\mathcal{P}_0$ and $\mathcal{P}_1$ is upper-bounded by*

$$p_{\mathrm{u}}(\mathcal{P}_0, \mathcal{P}_1) \leq 1 - \frac{1}{d^3}\left((d^2 - k)\frac{1}{\sqrt{d+1}} + k\right), \tag{5.16}$$

where k is the number of fixed points of the permutation π. Moreover, when $k = 0$, then

$$p_{\mathrm{u}}(\mathcal{P}_0, \mathcal{P}_1) \leq 1 - \frac{1}{d\sqrt{d+1}}. \tag{5.17}$$

Proof. Taking $\rho = \frac{\mathbf{1}}{d^2}$ we calculate

$$\begin{aligned}\sum_{i=1}^{d^2} |\langle x_i|\rho|y_i\rangle| &= \frac{1}{d^2}\sum_{i=1}^{d^2} |\langle x_i|y_i\rangle| = \frac{1}{d^3}\sum_{i=1}^{d^2} |\langle \phi_i|\phi_{\pi(i)}\rangle| \\ &= \frac{1}{d^3}\left((d^2-k)\frac{1}{\sqrt{d+1}} + k\right),\end{aligned} \tag{5.18}$$

where the last equality follows from the SIC condition in Eq. (3.16). ∎

5.3 Discrimination without entanglement

In this section, we will consider the special case of the problem of unambiguous discrimination of quantum measurements. We will assume that we cannot use the additional register, thus this scheme does not allow for the assistance of entanglement. We can only prepare an input state on a single register and apply the measurement contained in the black box to this state. Fortunately, the measurement gives a classical label, which can be used to decide which of the measurements was in the black box. This scheme is depicted in Figure 5.3. The post-processing, which, given the classical label from the black box measurement, allows us to make a final decision is denoted by a trivial final measurement $\mathcal{P}_F$.

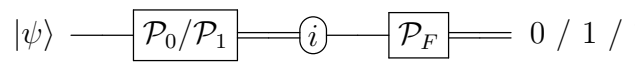

Figure 5.3: Scheme of discrimination of quantum measurements without the assistance of entanglement.

Our goal to is unambiguously discriminate measurements $\mathcal{P}_0$ and $\mathcal{P}_1$ without any additional register. We prepare an input state $|\psi\rangle \in \mathcal{D}(\mathcal{X})$ and apply the measurement in the black box. The probability of successful unambiguous discrimination without the use of entanglement yields

$$\tilde{p}_u(\mathcal{P}_F, \psi) = \frac{1}{2}\mathrm{tr}\left(\Omega_0\mathcal{P}_0(\psi)\right) + \frac{1}{2}\mathrm{tr}\left(\Omega_1\mathcal{P}_1(\psi)\right). \tag{5.19}$$

The unambiguity condition for the scheme without additional register can be written as

$$\mathrm{tr}\left(\Omega_0\mathcal{P}_1(\psi)\right) = \mathrm{tr}\left(\Omega_1\mathcal{P}_0(\psi)\right) = 0. \tag{5.20}$$

The optimized probability of unambiguous discrimination without the assistance of entanglement is given by

$$\tilde{p}_u \coloneqq \max_{\mathcal{P}_F,\psi} \tilde{p}_u(\mathcal{P}_F, \psi). \tag{5.21}$$

The following theorem concerns the discrimination of von Neumann measurements and states the optimal probability of their correct unambiguous discrimination without the assistance of entanglement [35].

Theorem 8 *The optimal success probability of unambiguous discrimination, without the use of entanglement, between von Neumann measurements $\mathcal{P}_{\mathbf{1}}$ and $\mathcal{P}_U$, is given by*

$$\tilde{p}_u = \frac{1}{2} \max_{A,B\in\{1,\ldots,d\}:A\cap B=\emptyset} \left\| \mathbb{P}_S \left(\sum_{i\in A} |i\rangle\langle i| + \sum_{j\in B} U|j\rangle\langle j|U^\dagger \right) \mathbb{P}_S \right\|, \tag{5.22}$$

where $\|\cdot\|$ is the operator norm and $\mathbb{P}_S$ in the orthogonal projection onto

$$S = \operatorname{span}\{U|i\rangle\}_{i\in\{1,\ldots,d\}\setminus A} \cap \operatorname{span}\{|j\rangle\}_{j\in\{1,\ldots,d\}\setminus B}. \tag{5.23}$$

Proof. Assume we have a fixed input state $\psi := |\psi\rangle\langle\psi|$. We will calculate the probability of unambiguous discrimination using the notation Δ for the dephasing channel. We will use the notation $\Omega_{\mathbf{1}}$ (Ω_U) to denote a measurement effects which corresponds to the situation that in the black box was $\mathcal{P}_{\mathbf{1}}$ ($\mathcal{P}_U$). We calculate

$$\begin{aligned}
\tilde{p_u}(\mathcal{P}_F,\psi) &= \frac{1}{2}\mathrm{tr}\left(\Omega_{\mathbf{1}}\mathcal{P}_{\mathbf{1}}(\psi)\right) + \frac{1}{2}\mathrm{tr}\left(\Omega_U\mathcal{P}_U(\psi)\right) \\
&= \frac{1}{2}\mathrm{tr}\left(\Omega_{\mathbf{1}}\sum_i |i\rangle\langle i|\psi|i\rangle\langle i|\right) + \frac{1}{2}\mathrm{tr}\left(\Omega_U\sum_j |j\rangle\langle j|U^\dagger\psi U|j\rangle\langle j|\right) \\
&= \frac{1}{2}\sum_i \mathrm{tr}\left(|i\rangle\langle i|\Omega_{\mathbf{1}}|i\rangle\langle i|\psi\right) + \frac{1}{2}\sum_j \mathrm{tr}\left(U|j\rangle\langle j|\Omega_U|j\rangle\langle j|U^\dagger\psi\right) \\
&= \frac{1}{2}\mathrm{tr}\left(\sum_i |i\rangle\langle i|\Omega_{\mathbf{1}}|i\rangle\langle i|\psi\right) + \frac{1}{2}\mathrm{tr}\left(U\sum_j |j\rangle\langle j|\Omega_U|j\rangle\langle j|U^\dagger\psi\right) \\
&= \frac{1}{2}\mathrm{tr}\left(\Delta\left(\Omega_{\mathbf{1}}\right)\psi\right) + \frac{1}{2}\mathrm{tr}\left(U\Delta\left(\Omega_U\right)U^\dagger\psi\right).
\end{aligned} \tag{5.24}$$

From the unambiguity condition we have

$$\begin{aligned}
\mathrm{tr}\left(\Omega_{\mathbf{1}}\mathcal{P}_U(\psi)\right) &= \mathrm{tr}\left(\Omega_{\mathbf{1}}\sum_j |j\rangle\langle j|U^\dagger\psi U|j\rangle\langle j|\right) \\
&= \mathrm{tr}\left(\sum_j U|j\rangle\langle j|\Omega_{\mathbf{1}}|j\rangle\langle j|U^\dagger\psi\right) \\
&= \mathrm{tr}\left(U\Delta\left(\Omega_{\mathbf{1}}\right)U^\dagger\psi\right) = 0
\end{aligned} \tag{5.25}$$

and

$$\begin{aligned}\mathrm{tr}\left(\Omega_U \mathcal{P}_{\mathbf{1}}(\psi)\right) &= \mathrm{tr}\left(\Omega_U \sum_i |i\rangle\langle i|\psi|i\rangle\langle i|\right) = \mathrm{tr}\left(\sum_i |i\rangle\langle i|\Omega_U|i\rangle\langle i|\psi\right) \\ &= \mathrm{tr}\left(\Delta\left(\Omega_U\right)\psi\right) = 0.\end{aligned} \tag{5.26}$$

In other words, $\psi \perp \mathrm{supp}\left(U\Delta\left(\Omega_{\mathbf{1}}\right)U^\dagger\right)$ and $\psi \perp \mathrm{supp}\left(\Delta\left(\Omega_U\right)\right)$. From the above we can see that we can restrict our attention to effects $\Omega_{\mathbf{1}}, \Omega_U$ which are diagonal. Moreover, the optimal effects can be chosen as projectors onto disjoint subsets A, B of $\{1, \ldots, d\}$. Hence

$$\begin{aligned}\tilde{p_u}(\mathcal{P}_F, \psi) &= \frac{1}{2}\mathrm{tr}\left(\Delta\left(\Omega_{\mathbf{1}}\right)\psi\right) + \frac{1}{2}\mathrm{tr}\left(U\Delta\left(\Omega_U\right)U^\dagger\psi\right) \\ &= \frac{1}{2}\mathrm{tr}\left(\Omega_{\mathbf{1}}\psi\right) + \frac{1}{2}\mathrm{tr}\left(U\Omega_U U^\dagger\psi\right) \\ &= \frac{1}{2}\mathrm{tr}\left(\sum_{i\in A}|i\rangle\langle i|\psi\right) + \frac{1}{2}\mathrm{tr}\left(U\left(\sum_{j\in B}|j\rangle\langle j|\right)U^\dagger\psi\right) \\ &= \frac{1}{2}\mathrm{tr}\left(\left(\sum_{i\in A}|i\rangle\langle i| + \sum_{j\in B}U|j\rangle\langle j|U^\dagger\right)\psi\right).\end{aligned} \tag{5.27}$$

When the disjoint subsets A and B are fixed, the maximum over input states ψ is, by linearity, equal to $\left\|\mathbb{P}_S\left(\sum_{i\in A}|i\rangle\langle i| + \sum_{j\in B}U|j\rangle\langle j|U^\dagger\right)\mathbb{P}_S\right\|$, where $\|\cdot\|$ is the operator norm. The orthogonal projectors $\mathbb{P}_S$ assure that the unambiguity condition is fulfilled. To obtain the probability of unambiguous discrimination one needs to optimize over disjoint subsets A, B of $\{1, \ldots, d\}$. ∎

5.4 Conditions for unambiguous discrimination

Similarly as in the symmetric discrimination discussed in the previous chapter, not all quantum measurements can be discriminated unambiguously. Sometimes we may need to use the black box many times, but it still may not be sufficient. The general conditions when quantum channels can be discriminated unambiguously with nonzero probability are formulated by the authors of [82]. We will state, as a proposition, a simplified version of these condition for the discrimination between two quantum channels. This proposition utilizes the notion of support of quantum channels, which was introduced in Eq. (2.61).

Proposition 9 ([82]) *If the quantum channels* $\{\Phi_0, \Phi_1\}$ *satisfy* $\mathrm{supp}(\Phi_0) \not\subseteq \mathrm{supp}(\Phi_1)$ *and* $\mathrm{supp}(\Phi_1) \not\subseteq \mathrm{supp}(\Phi_0)$*, then they can be unambiguously discriminated*

(with nonzero probability) by 2 uses. Otherwise, they cannot be unambiguously discriminated by any finite number of uses.

In the above proposition we do not specify the exact scheme of discrimination, but we only say whether unambiguous discrimination is possible or not. The detailed descriptions of the parallel and adaptive schemes for unambiguous discrimination will be presented in the following section.

In this chapter we focus on unambiguous discrimination of measurements with rank-one effects. Recall that for the measurement $\mathcal{P}$ with effects $\{|x_i\rangle\langle x_i|\}_i$, its Kraus operators have the form $\{|i\rangle\langle x_i|\}_i$. Therefore, we can draw a simple conclusion from the above Proposition, when quantum measurements with rank-one effects can be unambiguously discriminated with nonzero probability. This is formulated as the following corollary.

Corollary 8 *Let $\mathcal{P}_0, \mathcal{P}_1$ be quantum measurements with effects $\{|x_i\rangle\langle x_i|\}_i$ and $\{|y_i\rangle\langle y_i|\}_i$ respectively. If there exists i such that $|x_i\rangle$ and $|y_i\rangle$ are linearly independent, then $\mathcal{P}_0$ and $\mathcal{P}_1$ can be unambiguously discriminated (with nonzero probability) by 2 uses. Otherwise, they cannot be unambiguously discriminated by any finite number of uses.*

5.5 Multiple-shot unambiguous discrimination

Consider the scheme when the black box containing the measurement can be used N times in parallel. This scenario is depicted in Figure 5.4. The input state $|\psi\rangle$ is prepared on the compound register and we apply the black box measurement on first N registers. Therefore we obtain N classical labels and perform the final measurement on the remaining register. Basing on the label of the final measurement we make a decision whether in the black box there was either $\mathcal{P}_0$, $\mathcal{P}_1$, or the inconclusive answer.

Therefore, in the scheme of parallel unambiguous discrimination we actually need to discriminate channels $\mathcal{P}_0 \otimes \ldots \otimes \mathcal{P}_0$ and $\mathcal{P}_1 \otimes \ldots \otimes \mathcal{P}_1$ extended by the identity channel. In such case, the probability of unambiguous discrimination after N queries yields

$$p_u^{(N)} := \max_{\mathcal{P}_F, \psi} p_u^{(N)}(\mathcal{P}_F, \psi). \tag{5.28}$$

where

$$p_u^{(N)}(\mathcal{P}_F, \psi) = \frac{1}{2}\mathrm{tr}\left(\Omega_0\left(\Phi_0^{\otimes N} \otimes \mathbb{1}\right)(\psi)\right) + \frac{1}{2}\mathrm{tr}\left(\Omega_1\left(\Phi_1^{\otimes N} \otimes \mathbb{1}\right)(\psi)\right). \tag{5.29}$$

To obtain the upper bound on the success probability of unambiguous discrimination we can apply Theorem 7 for these channels. The resulting bound is stated

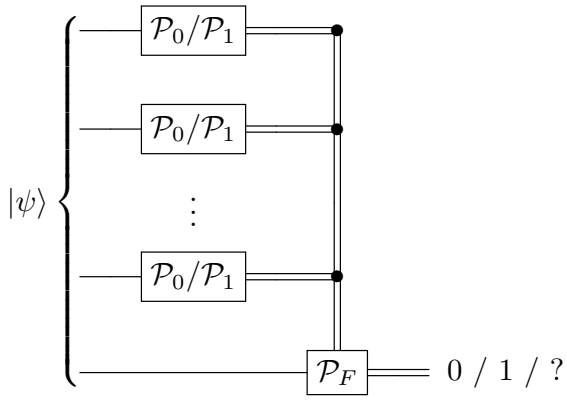

Figure 5.4: Scheme of unambiguous parallel discrimination of quantum measurements $\mathcal{P}_0$ and $\mathcal{P}_1$.

as the following Remark.

Remark 1 *The optimal success probability of unambiguous discrimination after N queries in the parallel scheme between rank-one measurements $\mathcal{P}_0$ and $\mathcal{P}_1$ with effects $\{|x_i\rangle\langle x_i|\}_{i=1}^m$ and $\{|y_i\rangle\langle y_i|\}_{i=1}^m$, respectively, is upper-bounded by*

$$p_{\mathrm{u}}^{(N)}(\mathcal{P}_0,\mathcal{P}_1) \leq 1 - \min_{\rho\in\mathcal{D}(\mathcal{X}^{\otimes N})} \sum_{i_1,\ldots,i_N} |\langle x_{i_1}\cdots x_{i_N}|\rho|y_{i_1}\cdots y_{i_N}\rangle|\,. \tag{5.30}$$

A similar bound in the case of von Neumann measurements is presented as the following Remark.

Remark 2 *The optimal success probability of unambiguous discrimination after N queries in the parallel scheme between von Neumann measurements $\mathcal{P}_{\mathbf{1}}$ and $\mathcal{P}_U$ is*

$$p_{\mathrm{u}}^{(N)}\left(\mathcal{P}_{\mathbf{1}^{\otimes N}},\mathcal{P}_{U^{\otimes N}}\right) = 1 - \min_{\rho\in\mathcal{D}(\mathcal{X}^{\otimes N})} \sum_i |\langle i|\rho U^{\otimes N}|i\rangle|. \tag{5.31}$$

Geometrical interpretation of the above Remark is presented in Figure 5.5.

The left figure represents the case of the first query to the black box, where we discriminate between von Neumann measurements $\mathcal{P}_U$ and $\mathcal{P}_{\mathbf{1}}$. The pair of the most distant eigenvalues of the optimized matrix UE_0 (see Eq. (3.15)) are denoted by $\lambda_1 = 1$ and λ_d, while the angle between them is denoted by Υ. The probability of unambiguous discrimination in this case corresponds to the distance from the point that is in the middle between λ_1 and λ_d to the unit circle. As we can see in the left picture, p_u is pretty small. Thus, to improve the discrimination we can use the black box two times in parallel. This is presented in the middle figure. Now, the probability of unambiguous discrimination $p_u^{(2)}$ is significantly bigger, but still

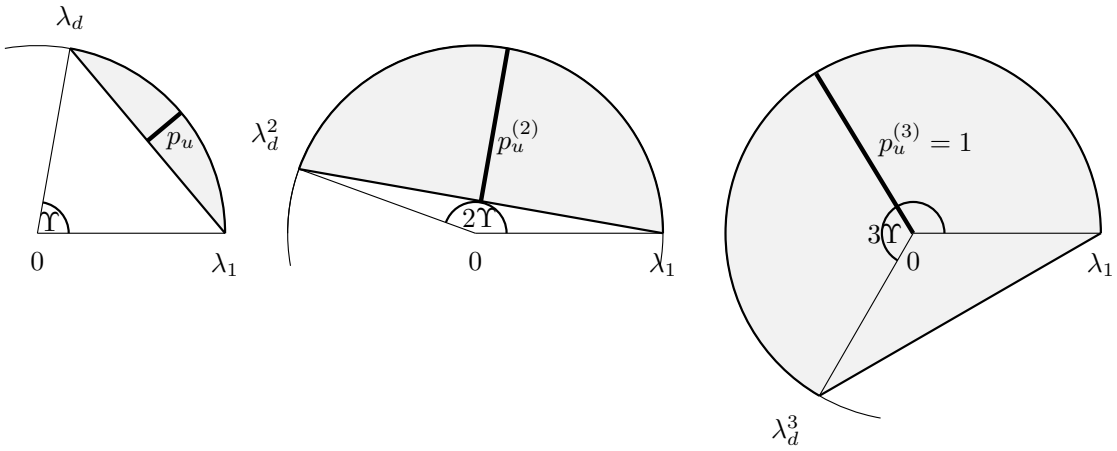

Figure 5.5: Geometrical interpretation of unambiguous discrimination of von Neumann measurements in the parallel scheme after first (left figure), second (middle figure) and third (right query) queries. The numerical range in each picture is contained in the gray area. p_u denoted the probability of unambiguous discrimination, λ_1 and λ_d denote the most distant eigenvalues of the unitary matrix and Υ stands for $\Upsilon(U)$.

smaller than one. Finally, after the third query to the black box, the numerical range contains zero. Thus, we obtain unambiguous discrimination, that is $p_u^{(3)} = 1$. This is depicted in the right figure.

Adaptive discrimination scheme allows us to perform processing between subsequent queries to the black box. We can use the obtained classical label to prepare an input for the next query. The scheme of adaptive unambiguous discrimination is depicted in Figure 5.6. In some cases, adaptive scheme may indeed lead to improvement in the discrimination, like it was in Section 4.5 for symmetric discrimination of SIC POVMs. However in the case of unambiguous discrimination of von Neumann measurements the use of adaptive scheme does not improve the discrimination.

Theorem 9 *Parallel scheme is optimal for unambiguous discrimination of von Neumann measurements.*

Proof. Without loss of generality we can assume that there were N queries to the black box and we discriminate between von Neumann measurements $\mathcal{P}_U$ and $\mathcal{P}_{\mathbf{1}}$. Moreover, we will assume that for the matrix U it holds that

$$\Upsilon(U) = \Theta(U), \tag{5.32}$$

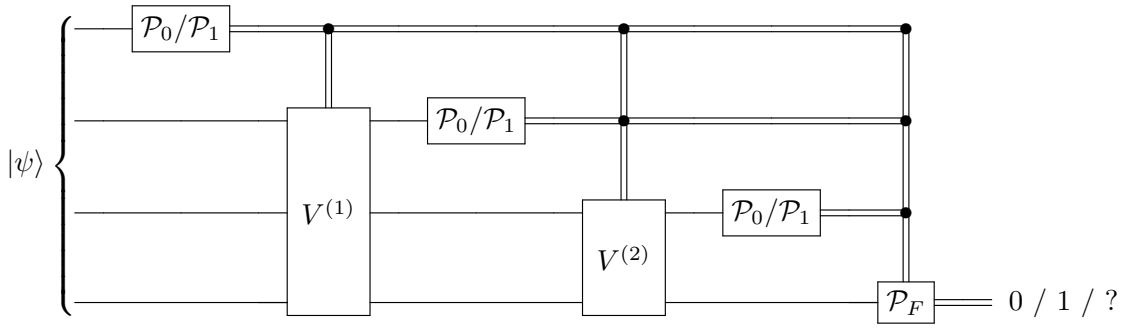

Figure 5.6: Scheme of adaptive unambiguous discrimination of quantum measurements $\mathcal{P}_0$ and $\mathcal{P}_1$.

where $\Upsilon(U)$ was defined in Eq. (3.15).

Let $|\psi\rangle$ be the input state. Let us denote

$$\begin{aligned} |x_i\rangle &= p_i^{-1/2} \left(\langle i| \otimes \mathbb{1}_{N+1}\right) A_{\mathbf{1}} |\psi_{A,B}\rangle \\ |y_i\rangle &= q_i^{-1/2} \left(\langle i| \otimes \mathbb{1}_{N+1}\right) A_U |\psi_{A,B}\rangle, \end{aligned} \tag{5.33}$$

where p_i and q_i are responsible for normalization and A_U and $A_{\mathbf{1}}$ are defined as

$$\begin{aligned} A_U = &\left(\mathbb{1}_{\mathcal{L}(\mathcal{X}_1,\ldots,\mathcal{X}_{N-1})} \otimes U \otimes \mathbb{1}_{\mathcal{L}(\mathcal{X}_{N+1})}\right) \\ &\times \left(\sum_{i_1,\ldots,i_{N-1}} |i_1 \cdots i_{N-1}\rangle\langle i_1 \cdots i_{N-1}| \otimes V^{(N-1)}_{i_1,\ldots,i_{N-1}} \right) \\ &\times \left(\mathbb{1}_{\mathcal{L}(\mathcal{X}_1,\ldots,\mathcal{X}_{N-2})} \otimes U \otimes \mathbb{1}_{\mathcal{L}(\mathcal{X}_N,\mathcal{X}_{N+1})}\right) \\ &\times \left(\sum_{i_1,\ldots,i_{N-2}} |i_1 \cdots i_{N-2}\rangle\langle i_1 \cdots i_{N-2}| \otimes V^{(N-2)}_{i_1,\ldots,i_{N-2}} \right) \\ &\cdots \\ &\times \left(\mathbb{1}_{\mathcal{L}(\mathcal{X}_1)} \otimes U \otimes \mathbb{1}_{\mathcal{L}(\mathcal{X}_3,\ldots,\mathcal{X}_{N+1})}\right) \\ &\times \left(\sum_{i_1} |i_1\rangle\langle i_1| \otimes V^{(1)}_{i_1} \right) \\ &\times \left(U \otimes \mathbb{1}_{\mathcal{L}(\mathcal{X}_2,\ldots,\mathcal{X}_{N+1})}\right). \end{aligned} \tag{5.34}$$

Repeating the calculation from the single-shot scenario from the proof of Theorem 7

we can upper-bound the probability of successful discrimination as follows

$$\begin{aligned} p_{\mathrm{u}}(\Psi_U, \Psi_{\mathbf{1}}) &\leq 1 - \min_{|\psi\rangle} \sum_i \left| \langle\psi| A_{\mathbf{1}}^\dagger \left(|i\rangle\langle i| \otimes \mathbb{1}_{N+1} \right) A_U |\psi\rangle \right| \\ &\leq 1 - \min_{|\psi\rangle} \left| \sum_i \langle\psi| A_{\mathbf{1}}^\dagger \left(|i\rangle\langle i| \otimes \mathbb{1}_{N+1} \right) A_U |\psi\rangle \right| \\ &= 1 - \min_{|\psi\rangle} \left| \langle\psi| A_{\mathbf{1}}^\dagger A_U |\psi\rangle \right|. \end{aligned} \tag{5.35}$$

From the work [100] we know that there exists a state $|\phi\rangle$ such that

$$|\langle\psi| A_{\mathbf{1}}^\dagger A_U |\psi\rangle| \geq |\langle\phi| U^{\otimes N} |\phi\rangle| \tag{5.36}$$

hods for all $|\psi\rangle$. Moreover, using optimality of U and Lemma 8, the state $|\phi\rangle$ can be chosen to satisfy $|\langle\phi|U^{\otimes N}|\phi\rangle| = \min_\rho \sum_i |\langle i|\rho U^{\otimes N}|i\rangle|$. This leads to the desired inequality

$$p_{\mathrm{u}}(\Psi_U, \Psi_{\mathbf{1}}) \leq 1 - \min_\rho \sum_i |\langle i|\rho U^{\otimes N}|i\rangle|. \tag{5.37}$$

■

Chapter 6

Asymmetric discrimination

In the previous chapters, we studied the discrimination between two quantum objects and wanted to know which of them was hidden in the black box. We either tried to minimize the chance of making a wrong decision in symmetric discrimination or to minimize the chance of getting the inconclusive result in unambiguous discrimination. Whenever we did not succeed in the discrimination process, we were saying that we failed and did not worry about what kind of mistake we had made.

In many real-life situations, objects which at first sight look similar have totally different values and prices. When discriminating such objects, it is a huge difference in what kind of mistake we make. When we hope to be given a cheap object, and by accident, we are given the expensive one, it would be a nice surprise. On the other hand, when we hope to be given the expensive object, and we are given the cheap one, that would be a bitter disappointment. This is the key idea of *asymmetric* discrimination. We differentiate between two types of errors and do not treat them equally. For example, we can assume some bounds on one type of error and study how small can the other type of error be. We can also study which of these types of errors can be equal to zero.

This approach to the discrimination problem is based on the statistical hypothesis testing. One of the quantum objects will be associated with the null hypothesis H_0 while the other object will be associated with the alternative hypothesis H_1. Intuitively, we can think that the object associated with the null hypothesis is the one we hope to be given. Using the language of black boxes, the null hypothesis states that in the black box, there is the object we hope to find there.

There may be many reasons why we hope that in the black box there is some specific object, and not the other one. When performing quantum computation, we apply quantum gates and measurements. The usefulness of the results of such computation depends on the quality of these gates and measurements. Hence, it is a legit question of how to verify if the given quantum gate or measurement acts exactly

the way it should. Such a verification process is known as *certification* [116, 117].

The term *certification* in general refers to verifying if some device works properly, that is, in a way we are promised it does. We will focus on the certification scheme which is based on binary hypothesis testing. The term *binary* in this context emphasizes that both null and alternative hypotheses are single-element sets. This can also be seen as certification of one quantum object *against* some other one. In the literature, the term certification may refer to a more general scheme [116, 117], e.g. when the alternative hypothesis corresponds to the set of objects, and this set can be infinite. Some of the results presented in this chapter were, in fact, originally proved also for this more general scheme, but in this dissertation, we will restrict our attention to the problem of binary certification. Therefore, in this chapter we will use the terms *certification* and *asymmetric discrimination* interchangeably.

This chapter will be based on the works [37] and [36]. We will begin by describing the problem of quantum hypothesis testing and introducing the schemes of certification for general quantum channels in Section 6.1. We will formulate the probabilities of making the false positive and false negative errors in single-shot and parallel discrimination schemes. In Section 6.2, we will prove a necessary and sufficient condition when a quantum channel can be certified against some other channel [37]. Similar conditions for the case of certification of quantum measurements will be presented in Section 6.3. Later, in Section 6.4 we will focus on the certification of SIC POVMs. Finally, certification of von Neumann measurements, covering the results from [36], will be explored in Section 6.5.

6.1 Quantum hypothesis testing

Let us now focus on the general problem of quantum hypothesis testing. We want to verify whether we were given channel the $\Phi_0 \in \mathcal{C}(\mathcal{X}, \mathcal{Y})$ or it was some other channel $\Phi_1 \in \mathcal{C}(\mathcal{X}, \mathcal{Y})$. We associate the channel Φ_0 with the H_0 hypothesis. If after the certification procedure we make a decision that we were given the channel Φ_0, then we *accept the hypothesis* H_0. Otherwise, we reject the null hypothesis and claim that the given channel was Φ_1, which is associated to the alternative hypothesis, H_1. Hence, our hypotheses can be succinctly written as

$$\begin{aligned} H_0 : &\quad \Phi_0; \\ H_1 : &\quad \Phi_1. \end{aligned} \tag{6.1}$$

While studying statistical hypothesis testing, we consider two types of errors. The type I error, known also as *false positive* error, happens when we reject the null hypothesis when in fact it was true. The converse situation, that is accepting the null hypothesis when the alternative one was true, is known as type II error or

false negative.

Single-shot certification

We will focus on the general, entanglement-assisted scheme. Similarly, as it was in the case of discrimination studied in the previous chapters, we can prepare a quantum state $|\psi\rangle$ on the compound space $\mathcal{X} \otimes \mathcal{Z}$ and apply the certified channel (either Φ_0 or Φ_1) on system $\mathcal{X}$. Hence, the possible output states are either

$$\rho_0^{|\psi\rangle} \coloneqq \left(\Phi_0 \otimes \mathbb{1}_{\mathcal{L}(\mathcal{Z})}\right)(|\psi\rangle\langle\psi|) \tag{6.2}$$

or

$$\rho_1^{|\psi\rangle} \coloneqq \left(\Phi_1 \otimes \mathbb{1}_{\mathcal{L}(\mathcal{Z})}\right)(|\psi\rangle\langle\psi|) \tag{6.3}$$

depending which of the channels was applied.

Having the output state either $\rho_0^{|\psi\rangle}$ or $\rho_1^{|\psi\rangle}$, we measure this state by the measurement $\mathcal{P}_F = \{\Omega, \mathbb{1} - \Omega\}$. If the measurement outcome corresponds to the effect Ω, then we accept the hypothesis H_0. Otherwise, that is when the measurement outcome corresponds to effect $\mathbb{1} - \Omega$, we reject the null hypothesis H_0 and accept the hypothesis H_1.

When the input state $|\psi\rangle$, and measurement effect Ω are fixed, then the probability of making the false positive error yields

$$\alpha(\psi, \Omega) \coloneqq \mathrm{Tr}\left((\mathbb{1} - \Omega)\,\rho_0^{|\psi\rangle}\right) = 1 - \mathrm{Tr}\left(\Omega \rho_0^{|\psi\rangle}\right), \tag{6.4}$$

and the probability of making the false negative error equals

$$\beta(\psi, \Omega) \coloneqq \mathrm{Tr}\left(\Omega \rho_1^{|\psi\rangle}\right). \tag{6.5}$$

The situation of special interest happens when we can be sure that we accept the null hypothesis only if it is indeed true. Therefore, we introduce the formal definition when a quantum channel *can be ϵ-certified* against some other channel.

Definition 4 *Let $\epsilon > 0$. Quantum channel Φ_0 can be ϵ-certified against channel Φ_1 if there exist an input state $|\psi\rangle$ and a measurement effect Ω such that $\beta(\psi, \Omega) = 0$ and $\alpha(\psi, \Omega) \leq \epsilon$.*

We will often use the notation for optimized probabilities of false positive error

$$\alpha_\delta \coloneqq \min_{\psi, \Omega} \{\alpha(\psi, \Omega) : \ \beta(\psi, \Omega) \leq \delta\} \tag{6.6}$$

and we will write

$$\alpha \coloneqq \alpha_0 = \min_{\psi,\Omega} \{\alpha(\psi,\Omega) : \ \beta(\psi,\Omega) = 0\}. \tag{6.7}$$

Parallel certification

The above scheme of certification can be directly extended to the parallel scheme. Assume that the certified channel can be used N times in the parallel setup. In other words, this time instead of certifying channel Φ_0 against Φ_1, we will be certifying channel $\Phi_0^{\otimes N}$ against the channel $\Phi_1^{\otimes N}$. As we will consider the general entanglement-assisted protocol, let $|\psi\rangle \in \mathcal{X}^{\otimes N} \otimes \mathcal{Z}$ be the input state. Hence, if the certified channel was Φ_0, then as the output state we have

$$\rho_0^{N,|\psi\rangle} \coloneqq \left(\Phi_0^{\otimes N} \otimes \mathbb{1}_{\mathcal{L}(\mathcal{Z})}\right)(|\psi\rangle\langle\psi|); \tag{6.8}$$

and analogously, if the certified channel was Φ_1, when we have the output state

$$\rho_1^{N,|\psi\rangle} \coloneqq \left(\Phi_1^{\otimes N} \otimes \mathbb{1}_{\mathcal{L}(\mathcal{Z})}\right)(|\psi\rangle\langle\psi|). \tag{6.9}$$

Now, the output state is in the larger space $\mathcal{Y}^{\otimes N} \otimes \mathcal{Z}$, hence we need to prepare a final measurement $\{\Omega, \mathbb{1} - \Omega\}$, where $\Omega \in \mathrm{Pos}\left(\mathcal{Y}^{\otimes N} \otimes \mathcal{Z}\right)$.

When the input state $|\psi\rangle$ and measurement effect Ω are fixed, then the probability of making the false positive error after N queries in the parallel scheme is given by

$$\alpha^{N,\mathbb{P}}(\psi,\Omega) \coloneqq \mathrm{Tr}\left((\mathbb{1} - \Omega)\,\rho_0^{N,|\psi\rangle}\right). \tag{6.10}$$

The probability of making the false negative error yields

$$\beta^{N,\mathbb{P}}(\psi,\Omega) \coloneqq \mathrm{Tr}\left(\Omega\rho_1^{N,|\psi\rangle}\right). \tag{6.11}$$

The resemblance between the notation used for the single-shot case and the parallel one is not incidental. In fact, when $N = 1$, then we recover the single-shot case, that is $\rho_0^{1,|\psi\rangle} = \rho_0^{|\psi\rangle}$ and $\rho_1^{1,|\psi\rangle} = \rho_1^{|\psi\rangle}$, as well as $\alpha^{1,\mathbb{P}}(\psi,\Omega) = \alpha(\psi,\Omega)$ and $\beta^{1,\mathbb{P}}(\psi,\Omega) = \beta(\psi,\Omega)$.

So far, the probabilities of making the false positive and false negative errors were introduced only for the case when the input state $|\psi\rangle$ and measurement effect Ω were fixed. However, in the task of certification we are allowed to prepare any input state as well as any final measurement, that is we can optimize over both. Therefore, let us introduce the optimized probability of making the false positive

error for parallel scheme as

$$\alpha_{\delta}^{N,\mathbb{P}} \coloneqq \min_{\psi,\Omega} \left\{ \alpha^{N,\mathbb{P}}(\psi,\Omega) : \ \beta^{N,\mathbb{P}}(\psi,\Omega) \leq \delta \right\} \tag{6.12}$$

and moreover, we introduce the notation

$$\alpha^{N,\mathbb{P}} \coloneqq \alpha_0^{N,\mathbb{P}}. \tag{6.13}$$

Finally, we will introduce the definition when we say that a quantum channel *can be certified* against some other channel.

Definition 5 *Quantum channel Φ_0 can be certified in the parallel scheme against channel Φ_1, if for every $\epsilon > 0$ there exist a natural number N, an input state $|\psi\rangle$ and a measurement effect Ω such that $\beta^{N,\mathbb{P}}(\psi,\Omega) = 0$ and $\alpha^{N,\mathbb{P}}(\psi,\Omega) \leq \epsilon$.*

6.2 Condition for certification of quantum channels

This section will be devoted to proving a theorem which gives necessary and sufficient condition when arbitrary quantum channel can be certified against some other quantum channel in a finite number of queries in the parallel scheme. This theorem was originally proved in [37].

Theorem 10 *Quantum channel Φ_0 can be certified against quantum channel Φ_1 in the parallel scheme in and inly if* $\operatorname{supp}(\Phi_0) \not\subseteq \operatorname{supp}(\Phi_1)$.

Before presenting the proof, we will introduce and prove two technical lemmas.

Lemma 2 *Let $\{|a_t\rangle\}_t$ and $\{|b_t\rangle\}_t$ be two orthonormal bases and $|\psi\rangle \coloneqq \sum_t \lambda_t |a_t\rangle|b_t\rangle$ where $\lambda_t > 0$ for every t. Let also $\rho_0^{|\psi\rangle}$ and $\rho_1^{|\psi\rangle}$ be as introduced in Eq.* (6.2),(6.3). *If* $\operatorname{supp}(\Phi_0) \not\subseteq \operatorname{supp}(\Phi_1)$, *then* $\operatorname{supp}\left(\rho_0^{|\psi\rangle}\right) \not\subseteq \operatorname{supp}\left(\rho_1^{|\psi\rangle}\right)$.

Proof of Lemma 2. Suppose by contradiction that $\operatorname{supp}\left(\rho_0^{|\psi\rangle}\right) \subseteq \operatorname{supp}\left(\rho_1^{|\psi\rangle}\right)$, that is

$$\operatorname{span}\left\{(E_i \otimes \mathbb{1})|\psi\rangle\right\}_i \subseteq \operatorname{span}\left\{(F_j \otimes \mathbb{1})|\psi\rangle\right\}_j. \tag{6.14}$$

Hence, for every i

$$(E_i \otimes \mathbb{1})|\psi\rangle = \sum_j \beta_j (F_j \otimes \mathbb{1})|\psi\rangle, \tag{6.15}$$

where β_j are not all equal to zero. As $|\psi\rangle \coloneqq \sum_t \lambda_t |a_t\rangle|b_t\rangle$, then we have

$$\begin{aligned}(E_i \otimes \mathbb{1})|\psi\rangle &= \sum_t \lambda_t \left(E_i|a_t\rangle \otimes |b_t\rangle\right); \\ \sum_j \beta_j (F_j \otimes \mathbb{1})|\psi\rangle &= \sum_t \lambda_t \sum_j \beta_j \left(F_j|a_t\rangle \otimes |b_t\rangle\right).\end{aligned} \tag{6.16}$$

As $\{|b_t\rangle\}_t$ is an orthonormal basis, then for every t

$$E_i|a_t\rangle = \sum_j \beta_j F_j |a_t\rangle, \tag{6.17}$$

and hence

$$E_i = \sum_j \beta_j F_j. \tag{6.18}$$

Therefore, $\text{span}\{E_i\}_i \subseteq \text{span}\{F_j\}_j$, which implies $\text{supp}(\Phi_0) \subseteq \text{supp}(\Phi_1)$. ∎

Lemma 3 *With the notation as above, if there exists a natural number N and an input state $|\psi\rangle$ such that* $\text{supp}\left(\rho_0^{N,|\psi\rangle}\right) \not\subseteq \text{supp}\left(\rho_1^{N,|\psi\rangle}\right)$, *then* $\text{supp}(\Phi_0) \not\subseteq \text{supp}(\Phi_1)$.

Proof of Lemma 3. Assume by contradiction that $\text{supp}(\Phi_0) \subseteq \text{supp}(\Phi_1)$, that is $\text{span}\left\{E_i\right\}_i \subseteq \text{span}\left\{F_j\right\}_j$. Hence it also holds that

$$\text{span}\left\{E_{i_1} \otimes \ldots \otimes E_{i_N}\right\}_{i_1,\ldots,i_N} \subseteq \text{span}\left\{F_{j_1} \otimes \ldots \otimes F_{j_N}\right\}_{j_1,\ldots,j_N}. \tag{6.19}$$

Thus, for every $i_1, \ldots, i_N$ we have that

$$E_{i_1} \otimes \ldots \otimes E_{i_N} = \sum_{j_1,\ldots,j_N} \beta_{j_1,\ldots,j_N} F_{j_1} \otimes \ldots \otimes F_{j_N}, \tag{6.20}$$

where not all $\beta_{j_1,\ldots,j_N}$ are equal to zero. Therefore, for every $i_1, \ldots, i_N$ it also holds that

$$\begin{aligned}\left(\left(E_{i_1} \otimes \ldots \otimes E_{i_N}\right) \otimes \mathbb{1}\right)|\psi\rangle &= \left(\left(\sum_{j_1,\ldots,j_N} \beta_{j_1,\ldots,j_N} F_{j_1} \otimes \ldots \otimes F_{j_N}\right) \otimes \mathbb{1}\right)|\psi\rangle \\ &= \sum_{j_1,\ldots,j_N} \beta_{j_1,\ldots,j_N} \left(\left(F_{j_1} \otimes \ldots \otimes F_{j_N}\right) \otimes \mathbb{1}\right)|\psi\rangle.\end{aligned} \tag{6.21}$$

This implies that

$$\begin{aligned}&\text{span}\left\{\left(\left(E_{i_1} \otimes \ldots \otimes E_{i_N}\right) \otimes \mathbb{1}\right)|\psi\rangle\right\}_{i_1,\ldots,i_N} \\ &\subseteq \text{span}\left\{\left(\left(F_{j_1} \otimes \ldots \otimes F_{j_N}\right) \otimes \mathbb{1}\right)|\psi\rangle\right\}_{j_1,\ldots,j_N}\end{aligned} \tag{6.22}$$

which contradicts with the assumption. Therefore, $\text{span}\{E_i\}_i \not\subseteq \text{span}\{F_j\}_j$, and eventually $\text{supp}(\Phi_0) \not\subseteq \text{supp}(\Phi_1)$. ∎

Now we are in position to present the proof of Theorem 10.

Proof of Theorem 10. ($\Longleftarrow$) Let $\text{supp}(\Phi_0) \not\subseteq \text{supp}(\Phi_1)$. From Lemma 2 this implies $\text{supp}\left(\rho_0^{|\psi\rangle}\right) \not\subseteq \text{supp}\left(\rho_1^{|\psi\rangle}\right)$, where the input state is $|\psi\rangle = \sum_t \lambda_t |a_t\rangle|b_t\rangle$. Hence, we can always find a state $|\phi\rangle$ for which

$$|\phi\rangle \not\perp \text{supp}\left(\rho_0^{|\psi\rangle}\right) \quad \text{and} \quad |\phi\rangle \perp \text{supp}\left(\rho_1^{|\psi\rangle}\right), \tag{6.23}$$

and therefore

$$\langle\phi|\rho_0^{|\psi\rangle}|\phi\rangle > 0 \quad \text{and} \quad \langle\phi|\rho_1^{|\psi\rangle}|\phi\rangle = 0. \tag{6.24}$$

Now we consider the certification scheme by taking the measurement with effects $\{\Omega, \mathbb{1} - \Omega\}$. Without loss of generality we can assume that $\Omega := |\phi\rangle\langle\phi|$ is a rank-one operator. Therefore

$$\text{tr}\left(\Omega\rho_0^{|\psi\rangle}\right) = \langle\phi|\rho_0^{|\psi\rangle}|\phi\rangle > 0, \tag{6.25}$$

and we calculate

$$\begin{aligned} \beta(\psi, \Omega) &= \text{tr}\left(\Omega\rho_1^{|\psi\rangle}\right) = \langle\phi|\rho_1^{|\psi\rangle}|\phi\rangle = 0; \\ \alpha(\psi, \Omega) &= \text{tr}\left((\mathbb{1} - \Omega)\rho_0^{|\psi\rangle}\right) = 1 - \langle\phi|\rho_0^{|\psi\rangle}|\phi\rangle < 1. \end{aligned} \tag{6.26}$$

Hence, after sufficiently many uses, N, of the certified channel in parallel (actually when $N \geq \lceil \frac{\log\epsilon}{\log\alpha} \rceil$) we obtain that $\text{tr}\left((\mathbb{1} - \Omega)^{\otimes N}\left(\rho_0^{|\psi\rangle}\right)^{\otimes N}\right) \leq \epsilon$ for any positive ϵ.Therefore, after N queries we will be able to exclude the false negative error.

($\Longrightarrow$) Assume that Φ_0 and Φ_1 can be certified in the parallel scenario. This means that there exist a natural number N, an input state $|\psi\rangle$ and a positive operator (measurement effect) Ω_0 on the composite system such that

$$\begin{aligned} \alpha^{N,\mathbb{P}}(\psi, \Omega) &= 1 - \text{tr}\left(\Omega\left(\Phi_0^{\otimes N} \otimes \mathbb{1}\right)(|\psi\rangle\langle\psi|)\right) \leq \epsilon < 1; \\ \beta^{N,\mathbb{P}}(\psi, \Omega) &= \text{tr}\left(\Omega\left(\Phi_1^{\otimes N} \otimes \mathbb{1}\right)(|\psi\rangle\langle\psi|)\right) = 0. \end{aligned} \tag{6.27}$$

Therefore $\text{tr}\left(\Omega\left(\Phi_0^{\otimes N} \otimes \mathbb{1}\right)(|\psi\rangle\langle\psi|)\right) > 0$ and thus

$$\begin{aligned} \Omega_0 &\not\perp \text{supp}\left(\left(\Phi_0^{\otimes N} \otimes \mathbb{1}\right)(|\psi\rangle\langle\psi|)\right) = \text{span}\left\{\left(E_{i_1} \otimes \ldots \otimes E_{i_N} \otimes \mathbb{1}\right)|\psi\rangle\right\}_{i_1,\ldots i_N}; \\ \Omega_0 &\perp \text{supp}\left(\left(\Phi_1^{\otimes N} \otimes \mathbb{1}\right)(|\psi\rangle\langle\psi|)\right) = \text{span}\left\{\left(F_{j_1} \otimes \ldots \otimes F_{j_N} \otimes \mathbb{1}\right)|\psi\rangle\right\}_{j_1,\ldots j_N}. \end{aligned} \tag{6.28}$$

Hence

$$\begin{aligned} &\text{span}\left\{(E_{i_1}\otimes\ldots\otimes E_{i_N}\otimes \mathbb{1})\,|\psi\rangle\right\}_{i_1,\ldots i_N} \\ &\not\subseteq \text{span}\left\{(F_{j_1}\otimes\ldots\otimes F_{j_N}\otimes \mathbb{1})\,|\psi\rangle\right\}_{j_1,\ldots j_N}. \end{aligned} \tag{6.29}$$

The reminder of the proof follows directly from Lemma 3. ■

From the above proof we have the following corollary which states what is the minimal number of queries needed to exclude false negative error.

Corollary 9 *The number of steps needed for parallel certification is bounded by*

$$N_\epsilon \geq \left\lceil \frac{\log \epsilon}{\log \alpha} \right\rceil, \tag{6.30}$$

where α is the upper bound on probability of making the false positive error in single-shot certification and ϵ is the upper bound on the probability of making the false positive error.

6.3 Conditions for certification of quantum measurements

In the previous section we proved a condition when arbitrary quantum channels can be certified in the parallel scheme. This condition was expressed in terms of Kraus operators of the given channels. In this section we will consider asymmetric discrimination of quantum measurements, and formulate conditions when they can be certified. Let us first focus on the most general class of quantum measurements.

Corollary 10 *Let $\mathcal{P}_0$ and $\mathcal{P}_1$ be quantum measurements with effects $\{M_i\}_{i=1}^m$ and $\{N_i\}_{i=1}^m$, respectively. Then, $\mathcal{P}_0$ can be certified against $\mathcal{P}_1$ in the parallel scheme if and only if there exists an index i, such that a pair of effects M_i, N_i satisfy* $\text{supp}(M_i) \not\subseteq \text{supp}(N_i)$.

Proof. Let

$$\begin{aligned} M_i &= \sum_{k_i} p^i_{k_i} |x^i_{k_i}\rangle\langle x^i_{k_i}| \\ N_i &= \sum_{l_i} q^i_{l_i} |x^i_{l_i}\rangle\langle x^i_{l_i}| \end{aligned} \tag{6.31}$$

be the spectral decompositions of M_i and N_i respectively (where $p^i_{k_i}, q^i_{l_i} > 0$ for every k_i, l_i). Then

$$\mathcal{P}_0(\rho) = \sum_i |i\rangle\langle i| \mathrm{tr}(\rho M_i) = \sum_i \sum_{k_i} p^i_{k_i} |i\rangle\langle x^i_{k_i}|\rho|x^i_{k_i}\rangle\langle i| \tag{6.32}$$

and hence the Kraus operators of $\mathcal{P}_0$ are $\left\{\sqrt{p^i_{k_i}}|i\rangle\langle x^i_{k_i}|\right\}_{i,k_i}$. Analogously,

$$\mathcal{P}_1(\rho) = \sum_i |i\rangle\langle i| \mathrm{tr}(\rho N_i) = \sum_i \sum_{l_i} q^i_{l_i} |i\rangle\langle y^i_{l_i}|\rho|y^i_{l_i}\rangle\langle i| \tag{6.33}$$

and hence the Kraus operators of $\mathcal{P}_1$ are $\left\{\sqrt{q^i_{l_i}}|i\rangle\langle y^i_{l_i}|\right\}_{i,l_i}$.

Therefore from Theorem 10 we have that $\mathcal{P}_0$ can be certified against $\mathcal{P}_1$ in the parallel scheme if and only if

$$\mathrm{span}\left\{\sqrt{p^i_{k_i}}|i\rangle\langle x^i_{k_i}|\right\}_{i,k_i} \not\subseteq \mathrm{span}\left\{\sqrt{q^i_{l_i}}|i\rangle\langle y^i_{l_i}|\right\}_{i,l_i}, \tag{6.34}$$

that is when there exists a pair of effects M_i, N_i for which $\mathrm{supp}(M_i) \not\subseteq \mathrm{supp}(N_i)$. ■

From the corollary concerning general quantum measurements we can formulate even simpler condition when measurements with rank-one effects can be certified in the parallel scheme.

Corollary 11 *Let $\mathcal{P}_0$ and $\mathcal{P}_1$ be quantum measurements with effects $\{p_i|x_i\rangle\langle x_i|\}_{i=1}^m$ and $\{q_i|y_i\rangle\langle y_i|\}_{i=1}^m$ for $p_i, q_i \in (0,1]$, respectively. Then, $\mathcal{P}_0$ can be certified against $\mathcal{P}_1$ in the parallel scheme if and only if there exists an index i, such that a pair of vectors $|x_i\rangle$, $|y_i\rangle$ is linearly independent.*

From the above corollary we can draw a conclusion when von Neumann measurements can be certified.

Corollary 12 *Von Neumann measurement $\mathcal{P}_U$ can be certified against von Neumann measurement $\mathcal{P}_V$ if and only if $U \neq V$.*

6.4 Certification of SIC POVMs

In this section, we will study certification of SIC POVMs of arbitrary dimension. Although it is an open question whether SIC POVMs exist in every dimension [62], in this section we will implicitly assume that they do exist. The following corollary,

which is a direct conclusion from Corollary 11, states when SIC POVMs can be certified.

Corollary 13 *SIC POVM $\mathcal{P}_0$ can be certified against SIC POVM $\mathcal{P}_1$ if and only if $\mathcal{P}_0 \neq \mathcal{P}_1$.*

Now, we will focus on the probability of making the false positive error when the false negative error cannot occur. We will present the results proved in [37]. First, the bound on the probability of making the false positive error for the single-shot case will be stated as Proposition 10. The bound for the parallel case will be formulated as Proposition 11.

Proposition 10 *Let $\mathcal{P}_0 = \{\frac{1}{d}|\phi_1\rangle\langle\phi_1|, \ldots, \frac{1}{d}|\phi_{d^2}\rangle\langle\phi_{d^2}|\}$ and $\mathcal{P}_1 = \{\frac{1}{d}|\psi_1\rangle\langle\psi_1|, \ldots, \frac{1}{d}|\psi_{d^2}\rangle\langle\psi_{d^2}|\}$ be SIC POVMs, where $|\psi_i\rangle = |\phi_{\pi(i)}\rangle$ for every $i = 1, \ldots, d^2$ and π is a permutation of d^2 elements with k fixed points. Assuming that the false negative error cannot occur, the probability of the obtaining the false positive error is upper bounded as follows*

$$\alpha \leq \frac{d+k}{d^2+d}. \tag{6.35}$$

Proof. To calculate the upper bound on α, we will calculate $\alpha(\psi, \Omega)$ for fixed input state ψ and final measurement Ω. As the input state we take the maximally entangled state $|\psi\rangle \coloneqq \frac{1}{\sqrt{d}}|\mathbb{1}\rangle\rangle$. If the measurement was $\mathcal{P}_0$, then the output state is

$$\rho_0^{|\psi\rangle} = (\mathcal{P}_0 \otimes \mathbb{1})(|\psi\rangle\langle\psi|) = \sum_{i=1}^{d^2} |i\rangle\langle i| \otimes \frac{1}{d^2}(|\phi_i\rangle\langle\phi_i|)^\top, \tag{6.36}$$

and similarly, if the measurement was $\mathcal{P}_1$, then the output state can be expressed as

$$\rho_1^{|\psi\rangle} = \sum_{i=1}^{d^2} |i\rangle\langle i| \otimes \frac{1}{d^2}(|\phi_{\pi(i)}\rangle\langle\phi_{\pi(i)}|)^\top. \tag{6.37}$$

Both output states have block-diagonal structure. Hence, without loss of generality we can restrict our attention to final measurements having block-diagonal structure, that is

$$\Omega \coloneqq \sum_{i=1}^{d^2} |i\rangle\langle i| \otimes \Omega_i^\top. \tag{6.38}$$

We are considering the situation when the false negative error cannot occur. To assure this condition, we need to make sure that $\Omega_i \perp |\phi_{\pi(i)}\rangle\langle\phi_{\pi(i)}|$ for every $i = 1, \ldots, d^2$. Therefore, we can take $\Omega_i \coloneqq \mathbb{1} - |\phi_{\pi(i)}\rangle\langle\phi_{\pi(i)}|$ for every $i = 1, \ldots, d^2$.

Having the input state and final measurement fixed, we will calculate the probability of obtaining the false positive error $\alpha(\psi,\Omega) = 1 - \mathrm{Tr}\left(\Omega\rho_0^{|\psi\rangle}\right)$. First we calculate

$$\begin{aligned}
\mathrm{tr}\left(\Omega\rho_0^{|\psi\rangle}\right) &= \mathrm{tr}\left(\left(\sum_{i=1}^{d^2}|i\rangle\langle i|\otimes\left(\mathbb{1}-|\phi_{\pi(i)}\rangle\langle\phi_{\pi(i)}|\right)^{\top}\right)\left(\sum_{j=1}^{d^2}|j\rangle\langle j|\otimes\frac{1}{d^2}(|\phi_j\rangle\langle\phi_j|)^{\top}\right)\right)\\
&= \frac{1}{d^2}\mathrm{tr}\left(\sum_{i=1}^{d^2}|i\rangle\langle i|\otimes\left(\mathbb{1}-|\phi_{\pi(i)}\rangle\langle\phi_{\pi(i)}|\right)^{\top}(|\phi_i\rangle\langle\phi_i|)^{\top}\right)\\
&= \frac{1}{d^2}\sum_{i=1}^{d^2}\langle\phi_i|\left(\mathbb{1}-|\phi_{\pi(i)}\rangle\langle\phi_{\pi(i)}|\right)|\phi_i\rangle = \frac{1}{d^2}\sum_{i=1}^{d^2}\left(1-|\langle\phi_i|\phi_{\pi(i)}\rangle|^2\right)\\
&= \frac{1}{d^2}\left(d^2-k\right)\left(1-|\langle\phi_i|\phi_{\pi(i)}\rangle|^2\right) = \frac{1}{d^2}\left(d^2-k\right)\left(1-\frac{1}{d+1}\right)\\
&= \frac{d^2-k}{d^2+d},
\end{aligned} \tag{6.39}$$

where k is the number of fixed points of the permutation π. Eventually, we can write the bound as

$$\alpha \leq \alpha\left(\psi,\Omega\right) = 1 - \mathrm{tr}(\Omega\rho_0^{|\psi\rangle}) = \frac{d+k}{d^2+d}. \tag{6.40}$$

■

Note that when the permutation π does not have fixed points, then $k = 0$, and the bound from the above proposition simplifies to $\alpha \leq \frac{1}{d+1}$. Similarly, if the permutation π has exactly one fixed point, then the above bound yields $\alpha \leq \frac{1}{d}$.

The above Proposition considered the certification of SIC POVMs in the single-shot scenario. In the remaining part of this section we will focus on the parallel certification scheme. The following proposition provides an upper bound on the probability of obtaining false positive error (when false negative error cannot occur) after N queries in parallel.

Proposition 11 *Let $\mathcal{P}_0 = \{\frac{1}{d}|\phi_1\rangle\langle\phi_1|,\ldots,\frac{1}{d}|\phi_{d^2}\rangle\langle\phi_{d^2}|\}$ and $\mathcal{P}_0 = \{\frac{1}{d}|\psi_1\rangle\langle\psi_1|,\ldots,\frac{1}{d}|\psi_{d^2}\rangle\langle\psi_{d^2}|\}$ be SIC POVMs, where $|\psi_i\rangle = |\phi_{\pi(i)}\rangle$ for every $i = 1,\ldots,d^2$ and π is a permutation of d^2 elements with k fixed points. Assuming that the false negative error cannot occur, the probability of the obtaining the false positive error after N queries in parallel is upper bounded as follows*

$$\alpha^{N,\mathbb{P}} \leq \left(\frac{d+k}{d^2+d}\right)^N. \tag{6.41}$$

Proof. This proof will be similar as the proof of Proposition 10. we will calculate the bound for particular choices of an input state and a final measurement. As for the input state, we will take the maximally entangled state. The output state of the certification procedure can be either

$$\begin{aligned}\rho_0^{N,|\psi\rangle} &= \left(\mathcal{P}_0^{\otimes N} \otimes \mathbb{1}\right)\left(\frac{1}{d^N}|\mathbb{1}\rangle\!\rangle\langle\!\langle\mathbb{1}|\right) \\ &= \frac{1}{d^N}\sum_{i_1,\ldots,i_N=1}^{d^2} |i_1\cdots i_N\rangle\langle i_1\cdots i_N| \otimes \frac{1}{d^N}\left(|\phi_{i_1}\cdots\phi_{i_N}\rangle\langle\phi_{i_1}\cdots\phi_{i_N}|\right)^\top \\ &= \frac{1}{d^{2N}}\sum_{i_1,\ldots,i_N=1}^{d^2} |i_1\cdots i_N\rangle\langle i_1\cdots i_N| \otimes \left(|\phi_{i_1}\cdots\phi_{i_N}\rangle\langle\phi_{i_1}\cdots\phi_{i_N}|\right)^\top\end{aligned} \tag{6.42}$$

if the measurement was $\mathcal{P}_0$, or

$$\rho_1^{N,|\psi\rangle} = \frac{1}{d^{2N}}\sum_{i_1,\ldots,i_N=1}^{d^2} |i_1\cdots i_N\rangle\langle i_1\cdots i_N| \otimes \left(|\phi_{\pi(i_1)}\cdots\phi_{\pi(i_N)}\rangle\langle\phi_{\pi(i_1)}\cdots\phi_{\pi(i_N)}|\right)^\top \tag{6.43}$$

if the measurement was $\mathcal{P}_1$. Similarly to the single-shot scenario, as both output states are block-diagonal, we can take the measurement effect with block-diagonal structure, which can be written as

$$\Omega \coloneqq \sum_{i_1,\ldots,i_N=1}^{d^2} |i_1\cdots i_N\rangle\langle i_1\cdots i_N| \otimes \Omega_{i_1\cdots i_N}^\top. \tag{6.44}$$

We require that the false negative error equals zero, which translates to $\Omega_{i_1\cdots i_N} \perp |\phi_{\pi(i_1)}\cdots\phi_{\pi(i_N)}\rangle\langle\phi_{\pi(i_1)}\cdots\phi_{\pi(i_N)}|$ for every multi-index $i_1\cdots i_N$. Therefore, we can define the blocks of Ω as

$$\Omega_{i_1\cdots i_N} \coloneqq \mathbb{1} - |\phi_{\pi(i_1)}\cdots\phi_{\pi(i_N)}\rangle\langle\phi_{\pi(i_1)}\cdots\phi_{\pi(i_N)}|. \tag{6.45}$$

We calculate

$$\begin{aligned}
\mathrm{tr}\left(\Omega\rho_0^{N,|\psi\rangle}\right) &= \mathrm{tr}\left(\left(\sum_{i_1,\ldots,i_N=1}^{d^2} |i_1\cdots i_N\rangle\langle i_1\cdots i_N| \otimes \left(\mathbb{1} - |\phi_{\pi(i_1)}\cdots\phi_{\pi(i_N)}\rangle\langle\phi_{\pi(i_1)}\cdots\phi_{\pi(i_N)}|\right)^\top\right)\right. \\
&\qquad \left.\left(\frac{1}{d^{2N}}\sum_{j_1,\ldots,j_N=1}^{d^2} |j_1\cdots j_N\rangle\langle j_1\cdots j_N| \otimes \left(|\phi_{j_1}\cdots\phi_{j_N}\rangle\langle\phi_{j_1}\cdots\phi_{j_N}|\right)^\top\right)\right) \\
&= \frac{1}{d^{2N}}\mathrm{tr}\left(\sum_{i_1,\ldots,i_N=1}^{d^2}\sum_{j_1,\ldots,j_N=1}^{d^2} |i_1\cdots i_N\rangle\langle i_1\cdots i_N|j_1\cdots j_N\rangle\langle j_1\cdots j_N|\right. \\
&\qquad \left.\otimes \left(\mathbb{1} - |\phi_{\pi(i_1)}\cdots\phi_{\pi(i_N)}\rangle\langle\phi_{\pi(i_1)}\cdots\phi_{\pi(i_N)}|\right)^\top \left(|\phi_{j_1}\cdots\phi_{j_N}\rangle\langle\phi_{j_1}\cdots\phi_{j_N}|\right)^\top\right) \\
&= \frac{1}{d^{2N}}\sum_{i_1,\ldots,i_N=1}^{d^2} \mathrm{tr}\left(\left(\mathbb{1} - |\phi_{\pi(i_1)}\cdots\phi_{\pi(i_N)}\rangle\langle\phi_{\pi(i_1)}\cdots\phi_{\pi(i_N)}|\right)|\phi_{i_1}\cdots\phi_{i_N}\rangle\langle\phi_{i_1}\cdots\phi_{i_N}|\right) \\
&= \frac{1}{d^{2N}}\sum_{i_1,\ldots,i_N=1}^{d^2} \left(1 - \left|\langle\phi_{i_1}\cdots\phi_{i_N}|\phi_{\pi(i_1)}\cdots\phi_{\pi(i_N)}\rangle\right|^2\right) \\
&= 1 - \frac{1}{d^{2N}}\sum_{i_1,\ldots,i_N=1}^{d^2} \left|\langle\phi_{i_1}\cdots\phi_{i_N}|\phi_{\pi(i_1)}\cdots\phi_{\pi(i_N)}\rangle\right|^2.
\end{aligned} \tag{6.46}$$

Therefore, we have

$$\alpha^{N,\mathbb{P}}(\psi,\Omega) = 1 - \mathrm{tr}\left(\Omega\rho_0^{N,|\psi\rangle}\right) = \frac{1}{d^{2N}}\sum_{i_1,\ldots,i_N=1}^{d^2} \left|\langle\phi_{i_1}\cdots\phi_{i_N}|\phi_{\pi(i_1)}\cdots\phi_{\pi(i_N)}\rangle\right|^2. \tag{6.47}$$

To get the exact upper bound we need to calculate the sum, that is to explain that

$$\sum_{i_1,\ldots,i_N=1}^{d^2} \left|\langle\phi_{i_1}\cdots\phi_{i_N}|\phi_{\pi(i_1)}\cdots\phi_{\pi(i_N)}\rangle\right|^2 = \sum_{s=0}^{N}\binom{N}{N-s}k^{N-s}(d^2-k)^s\frac{1}{(d+1)^s}. \tag{6.48}$$

First, note that

$$\left|\langle\phi_{i_1}\cdots\phi_{i_N}|\phi_{\pi(i_1)}\cdots\phi_{\pi(i_N)}\rangle\right|^2 = \left|\langle\phi_{i_1}|\phi_{\pi(i_1)}\rangle\cdots\langle\phi_{i_N}|\phi_{\pi(i_N)}\rangle\right|^2 = \frac{1}{(d+1)^s}, \tag{6.49}$$

where $s \coloneqq |\{i_l : \ i_l \neq \pi(i_l)\}|$. In other words, every time we encounter a fixed point of the permutation we have a factor $\langle\phi_{i_j}|\phi_{\pi(i_j)}\rangle$ which is equal to one. Let

us now focus on consecutive factors of the right hand site of the Eq. (6.48). The factor $\binom{N}{N-s}$ corresponds to choosing $N-s$ elements for which $\langle\phi_{i_j}|\phi_{\pi(i_j)}\rangle = 1$. Then, on each of those elements there can be one of k elements (as k stands for the number of fixed points of the permutation π), therefore k^{N-s}. Then, on the remaining s elements there can one of d^2-k values which are not fixed points of the permutation, hence we obtain $(d^2-k)^s$. Further calculations reveal the concise expression for the upper bound on the probability of making the false negative error, that is

$$\begin{aligned}\alpha^{N,\mathbb{P}} \leq \alpha^{N,\mathbb{P}}(\psi,\Omega) &= \frac{1}{d^{2N}}\sum_{s=0}^{N}\binom{N}{N-s}k^{N-s}(d^2-k)^s\frac{1}{(d+1)^s} \\ &= \left(\frac{d+k}{d^2+d}\right)^N.\end{aligned} \tag{6.50}$$

■

In the case of permutation π having no fixed points, that is when $k=0$, the above bound simplifies to $\alpha^{N,\mathbb{P}} \leq \left(\frac{1}{d+1}\right)^N$.

6.5 Certification of von Neumann measurements

In this section, we will study the problem of asymmetric discrimination from a bit different perspective. We will no longer try to exclude the false negative error, but focus on how small it can be, even when zero cannot be achieved. More precisely, we will assume an upper bound on the probability of making the false positive error. This bound is known in statistics as *statistical significance* and it will be denoted by δ. Our goal will be to calculate how small can be the probability of making the false negative error assuming the statistical significance δ.

Let us introduce the notation. First, recall that the probability of making the false negative error was denoted $\beta(\psi,\Omega) \coloneqq \operatorname{tr}\left(\Omega\rho_1^{|\psi\rangle}\right)$, where $\rho_1^{|\psi\rangle}$ was defined in Eq. (6.9). This probability depends on the input state ψ and the measurement $\{\Omega, \mathbb{1}-\Omega\}$. We define

$$\beta_\delta \coloneqq \min_{\psi,\Omega}\{\beta(\psi,\Omega) : \alpha(\psi,\Omega) \leq \delta\} \tag{6.51}$$

as the optimized probability of making the false negative error when the probability of making the false positive error is no greater than δ.

The following theorem, proved in [36], provides an expression for this optimal probability of making the false negative error with statistical significance δ. The technical proof is postponed to Appendix B.

Theorem 11 *Consider the problem of certification of von Neumann measurements where $\mathcal{P}_{\mathbf{1}}$ corresponds to the H_0 hypothesis and $\mathcal{P}_U$ corresponds to H_1 hypothesis. Then*

$$\beta_\delta = \max_{E \in \mathcal{DU}(\mathcal{X})} \nu^2_{\sqrt{1-\delta}}(UE)\,. \tag{6.52}$$

The geometrical representation of the result from Theorem 11 is depicted in Figure 6.1. The notation in this figure is very similar to the notation used in Figures 3.4 and 5.2. The pair of the most distance eigenvalue of the optimized unitary matrix UE_0 is denoted by λ_1 and λ_d. They are represented on the unit circle on the complex plane. The expression in Theorem 11 can be seen as square of the distance between zero and the set q-numerical range, where $q = \sqrt{1-\delta}$. In the picture, the set q-numerical range in contained in the gray area. The detailed shape of this set can be found in [79]. The distance between zero and q-numerical range is denoted by ν. We can see that it depends not only on the eigenvalues of the unitary matrix, but also on the parameter δ. This coincides with the intuition that the bigger can be the upper bound on the false positive error, the smaller gets the optimal probability of making the false negative error.

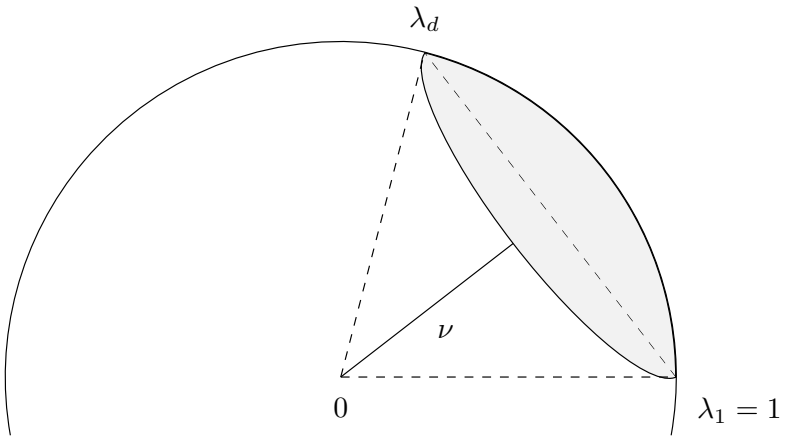

Figure 6.1: Geometrical interpretation of the probability of making the false negative error in the asymmetric discrimination of von Neumann measurements. λ_1 and λ_d denote the most distant eigenvalue of a matrix UE_0, where E_0 is the optimal diagonal unitary matrix from Eq. (6.52). The q-numerical range is contained in the gray area and ν denotes the distance from zero to the q-numerical range.

Parallel certification

Similarly as for symmetric and unambiguous discrimination, we will study parallel discrimination of von Neumann measurements. Consider the situation when the

certified measurement can be used N times in parallel. The scheme of parallel certification is analogous as the ones described for symmetric and unambiguous discrimination.

We assume that, after N queries, the probability of making the false positive error is no greater than δ. The optimal probability of making the negative error after N queries in parallel yields

$$\beta_{\delta}^{N,\mathbb{P}} \coloneqq \min_{\psi,\Omega} \left\{ \beta^{N,\mathbb{P}}(\psi,\Omega) : \alpha^{N,\mathbb{P}}(\psi,\Omega) \leq \delta \right\}. \tag{6.53}$$

The following corollary gives us the expression for $\beta_{\delta}^{N,\mathbb{P}}$.

Corollary 14 *Consider the problem of parallel certification of von Neumann measurements where $\mathcal{P}_{\mathbf{1}}$ corresponds to the H_0 hypothesis and $\mathcal{P}_U$ corresponds to H_1 hypothesis. Assume that there were N queries in the parallel scheme. Then*

$$\beta_{\delta}^{N,\mathbb{P}} = \max_{E \in \mathcal{DU}(\mathcal{X})} \nu^2_{\sqrt{1-\delta}} \left(U^{\otimes N} E^{\otimes N} \right). \tag{6.54}$$

Proof. We will apply Theorem 11 for von Neumann measurements $\mathcal{P}_{\mathbf{1}^{\otimes N}}$ and $\mathcal{P}_{U^{\otimes N}}$. We have

$$\beta_{\delta}^{N,\mathbb{P}} = \max_{E \in \mathcal{DU}(\mathcal{X}^{\otimes N})} \nu^2_{\sqrt{1-\delta}} \left(U^{\otimes N} E \right). \tag{6.55}$$

From Theorem 5 we have that

$$\max_{E \in \mathcal{DU}(\mathcal{X}^{\otimes N})} \nu^2_{\sqrt{1-\delta}} \left(U^{\otimes N} E \right) = \max_{E \in \mathcal{DU}(\mathcal{X})} \nu^2_{\sqrt{1-\delta}} \left(U^{\otimes N} E^{\otimes N} \right). \tag{6.56}$$

■

The following Figure 6.2 represents the greometrical interpretation of the parallel discrimination. The left picture corresponds to the single-shot case in Fig 6.1. The right picture corresponds to two queries in the parallel scheme. The most distant eigenvalues are this time $\lambda_1 = 1$ and λ_d^2. The q-numerical range is contained in the gray area and we can see that it contains zero. Therefore, assuming the bound δ on the probability of making the false positive error, after two queries in the parallel scheme we can exclude the false negative error. Note that in this case, the numerical range does not contain zero, as the closest point from numerical range to zero in the point in the middle between $\lambda_1 = 1$ and λ_d^2. Thus, perfect discrimination in the symmetric scheme would not be possible.

Can the use of processing between subsequent queries to the black box help in the case of asymmetric discrimination? For the discrimination of von Neumann measurements in both symmetric and unambiguous schemes the parallel scheme was optimal. On the other hand, symmetric discrimination of SIC POVMs could be improved by the use of adaptive scheme. The following theorem states

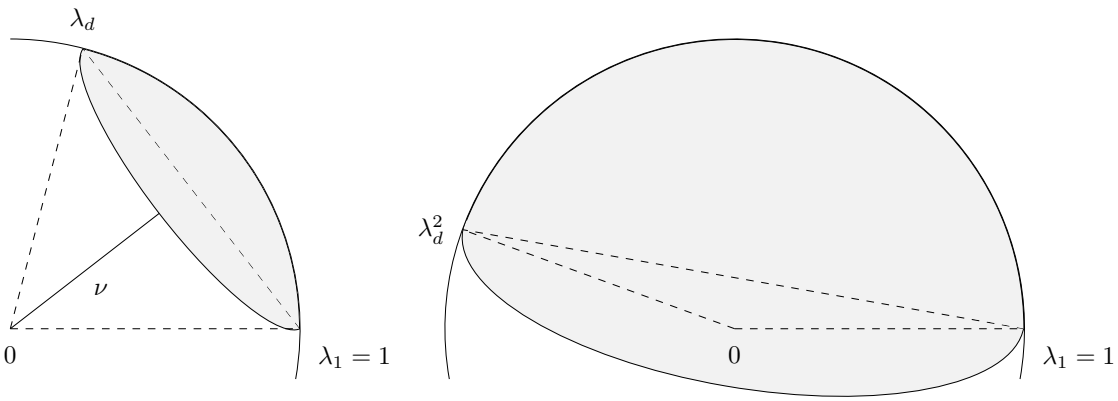

Figure 6.2: Geometrical interpretation of the probability making the false negative error in the asymmetric discrimination between von Neumann measurements in the single-shot scenario (left figure) and after two queries in parallel (right figure). λ_1 and λ_d denote the most distant eigenvalue of a matrix UE_0, where E_0 is the optimal diagonal unitary matrix from Eq. (6.52). The q-numerical ranges are contained in the gray areas. The distance from zero to the q-numerical range is denoted by ν. In the right figure, zero is contained in the q-numerical range, that is $\nu = 0$.

that for asymmetric discrimination of von Neumann measurements, the parallel discrimination scheme is optimal.

Theorem 12 *The parallel scenario for certification of von Neumann measurements is optimal, i.e. the optimal probability of making the false negative error in the adaptive scheme cannot be smaller than in the parallel scheme.*

Proof. We will present the detailed proof for the case when the number of queries to the black box equals $N = 3$. The proof for any greater number of queries is analogous and it can be found in [36].

Assume we are discriminating von Neumann measurements $\mathcal{P}_{\mathbf{1}}, \mathcal{P}_U$ of dimension d. We will use the notation $\Xi^{(N)}(\cdot)$ to denote an adaptive discrimination scheme which takes N copies of the input (in our case $N = 3$) and outputs a quantum state. Let $d_1 \leq d_2 \leq d_3$ be natural numbers and $d_3 = d_3' d_3''$. These numbers will denote dimensions of auxiliary systems in the construction of $\Xi^{(3)}$. This adaptive scheme is depicted in Figure 6.3, where we use the notation $\mathcal{P}_?$ for a black box containing either the measurement $\mathcal{P}_{\mathbf{1}}$ or $\mathcal{P}_U$. More formally, this scheme can be written as a composition of quantum operations as

$$\begin{aligned}\Xi^{(3)}\left(\mathcal{P}_?\right) = \left(\mathbb{1}_{d^2} \otimes \mathcal{P}_? \otimes \mathbb{1}_{d_3'} \otimes \mathrm{tr}_{d_3''}\right) \circ \Xi_2 \circ \left(\mathbb{1}_d \otimes \mathcal{P}_? \otimes \mathbb{1}_d \otimes \mathbb{1}_{d_2}\right) \\ \circ \Xi_1 \circ \left(\mathcal{P}_? \otimes \mathbb{1}_{d^2} \otimes \mathbb{1}_{d_1}\right)\left(|\psi\rangle\langle\psi|\right),\end{aligned} \tag{6.57}$$

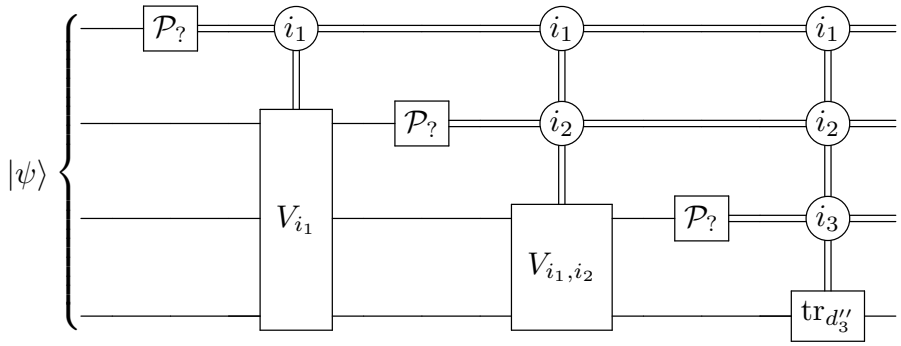

Figure 6.3: Scheme of the adaptive discrimination of quantum measurements from Eq. (6.57).

where $\Xi_j(X) = W_j X W_j^\dagger$ for $j = 1, 2$ and

$$\begin{aligned} W_1 &= \sum_{i_1} |i_1\rangle\langle i_1| \otimes V_{i_1}; \\ W_2 &= \sum_{i_1,i_2} |i_1 i_2\rangle\langle i_1 i_2| \otimes V_{i_1,i_2}. \end{aligned} \tag{6.58}$$

Now, we will use the fact that every von Neumann measurement $\mathcal{P}_U$ can be written as $\mathcal{P}_U = \Delta \circ \Phi_{(UE_0)^\dagger}$, where E_0 maximizes the expression in Eq. (6.54) in Corollary 14. Hence, the above adaptive scheme can be rewritten as

$$\Xi^{(3)}(\mathcal{P}_?) = \left(\Delta^{\otimes 3} \otimes \mathbb{1}_{d_3'} \otimes \mathrm{tr}_{d_3''}\right) \circ \Xi^{(3)}\left(\Phi_?\right), \tag{6.59}$$

where

$$\begin{aligned} \Xi^{(3)}\left(\Phi_?\right) = \left(\mathbb{1}_{d^2} \otimes \Phi_? \otimes \mathbb{1}_{d_3}\right) \circ \Xi_2 \circ \left(\mathbb{1}_d \otimes \Phi_? \otimes \mathbb{1}_d \otimes \mathbb{1}_{d_2}\right) \\ \circ \Xi_1 \circ \left(\Phi_? \otimes \mathbb{1}_{d^2} \otimes \mathbb{1}_{d_1}\right)\left(|\psi\rangle\langle\psi|\right) \end{aligned} \tag{6.60}$$

and $\Phi_? \in \left\{\mathcal{P}_{\mathbb{1}}, \mathcal{P}_{(UE_0)^\dagger}\right\}$. This version of the adaptive discrimination scheme is depicted in Figure 6.4.

Now we note that $\Xi^{(3)}\left(\Phi_{\mathbb{1}}\right)$ and $\Xi^{(3)}\left(\Phi_{(UE_0)^\dagger}\right)$ are actually the adaptive schemes of discrimination of unitary channels $\Phi_{\mathbb{1}}$ and $\Phi_{(UE_0)^\dagger}$. In such a scheme, using the Data Processing Inequality in Lemma 9 in Appendix B, the probability of making the false negative error in asymmetric discrimination of $\mathcal{P}_{\mathbb{1}}$ and $\mathcal{P}_U$ cannot be smaller than the probability of making the false negative error in asymmetric discrimination of $\Phi_{\mathbb{1}}$ and $\Phi_{(UE_0)^\dagger}$.

From the work [118] we know that the minimized probability of making the false negative error for N copier of unitary channels is achieved in the parallel scheme. Let

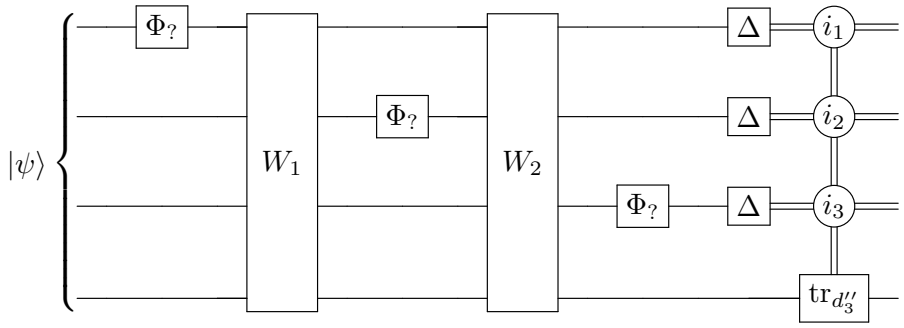

Figure 6.4: Scheme of the adaptive discrimination of unitary channels from Eq. (6.60).

$\tilde{\beta}_\delta^{N,\mathbb{P}}$ be the minimized probability of making the false negative error for asymmetric discrimination of states $\Xi^{(N)}(\mathcal{P}_{\mathbb{1}})$ and $\Xi^{(N)}(\mathcal{P}_U)$. From Proposition 14 we can see that the optimal probability of making the false negative error for asymmetric discrimination of unitary channels $\Phi_{\mathbb{1}}$ and $\Phi_{(UE_0)^\dagger}$ equals $\nu^2_{\sqrt{1-\delta}}\left(U^{\otimes N}E_0^{\otimes N}\right)$, thus

$$\tilde{\beta}_\delta^{N,\mathbb{P}} \geq \nu^2_{\sqrt{1-\delta}}\left(U^{\otimes N}E_0^{\otimes N}\right). \tag{6.61}$$

From Corollary 14 we know that

$$\beta_\delta^{N,\mathbb{P}} = \nu^2_{\sqrt{1-\delta}}\left(U^{\otimes N}E_0^{\otimes N}\right), \tag{6.62}$$

hence eventually we obtain

$$\tilde{\beta}_\delta^{N,\mathbb{P}} \geq \beta_\delta^{N,\mathbb{P}}. \tag{6.63}$$

■

Chapter 7

Conclusions

This dissertation studied the problem of discrimination of quantum measurements in symmetric, unambiguous and asymmetric settings. We proved numerous results concerning probabilities of successful discrimination in single-shot and multiple-shot scenarios. We studied when adaptive schemes can be helpful for discrimination.

As for the symmetric discrimination, we first analyzed the naïve discrimination scheme when we only prepared an input state and measured it with a measurement hidden in the black box. We calculated the probability of successful discrimination in this scheme. Later, we considered the general entanglement-assisted discrimination scheme. We proved a condition when von Neumann measurements could be discriminated perfectly in the single-shot scenario. We also calculated the diamond norm distance between von Neumann measurements which allowed us to calculate the optimal probability of discrimination between von Neumann measurements and present the geometrical interpretation of this quantity. We also studied discrimination of other measurements with rank-one effects and focused on SIC POVMs. We calculated the lower and upper bounds on the diamond norm distance between SIC POVMs and addressed the problem when the use of entangled input improves the discrimination. We characterized the discrimination of qubit SIC POVMs and studied the chance of their discrimination in the asymptotic limit, when the dimension tends to infinity.

We introduced the parallel and adaptive discrimination schemes and calculated the probability of successful discrimination between von Neumann measurements after N queries in parallel. We also calculated the minimal number of queries needed to discriminate von Neumann measurements in the parallel scheme perfectly. Interestingly, it turned out that for the discrimination of von Neumann measurement, the parallel scheme is optimal in the sense that the use of an adaptive scheme cannot improve the probability of successful discrimination. We calculated the diamond norm distance between tensor products of SIC POVMs and characterized when the SIC POVM of dimension two can be discriminated perfectly in the multiple-shot

case. It turned out that for the qubit case, they could be either discriminated perfectly after two queries in the parallel scheme, or they cannot be discriminated perfectly after any finite number of queries.

Moreover, we formulated conditions when a pair of quantum measurements require an adaptive scheme to be discriminated perfectly, that is when perfect discrimination cannot be achieved in the parallel scheme but it is achieved in the adaptive scheme. We presented an example of a pair of SIC POVMs of dimension three which cannot be discriminated perfectly after any finite number of queries in the parallel scheme, but can be discriminated perfectly after two queries in the adaptive scheme. We described the detailed algorithm of such an adaptive discrimination scheme.

As far as unambiguous discrimination is concerned, we calculated the probability of unambiguous discrimination between two measurements having rank-one effects both in the single-shot and parallel cases. We presented the geometrical interpretation of this probability for discrimination of von Neumann measurements and considered the particular case of unambiguous discrimination without the assistance of entanglement. We found an expression for the probability of unambiguous discrimination of SIC POVMs. We also proved that for the unambiguous discrimination of von Neumann measurements, the use of an adaptive scheme could not improve the discrimination; that is, the parallel scheme is optimal.

We also studied asymmetric discrimination, which was based on hypothesis testing. We proved a condition when general quantum channels could be discriminated in the asymmetric scheme; that is when we can assure that no false negative error can occur. We formulated similar conditions for general quantum measurements as well as von Neumann measurements and SIC POVMs. We calculated the optimal probability of making the false positive error for SIC POVMs for single-shot and parallel schemes.

Furthermore, we considered the case when we assumed an upper bound on the false positive error and wanted to find the optimal probability of making the false negative error. We calculated this optimal probability for von Neumann measurements and presented its geometrical interpretation and connection with the notion of q-numerical range. We also analyzed the multiple-shot discrimination and proved that the adaptive scheme could not improve the asymmetric discrimination of von Neumann measurements.

Bibliography

[1] S. Russell and P. Norvig, "Artificial Intelligence: A Modern Approach," 2002.

[2] A. Likas, N. Vlassis, and J. J. Verbeek, "The global k-means clustering algorithm," *Pattern Recognition*, vol. 36, no. 2, pp. 451–461, 2003.

[3] J. A. Hartigan and M. A. Wong, "Algorithm as 136: A k-means clustering algorithm," *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, vol. 28, no. 1, pp. 100–108, 1979.

[4] K. Wagstaff, C. Cardie, S. Rogers, S. Schrödl, *et al.*, "Constrained k-means clustering with background knowledge," in *ICML*, vol. 1, pp. 577–584, 2001.

[5] F. Angiulli and C. Pizzuti, "Fast outlier detection in high dimensional spaces," in *European Conference on Principles of Data Mining and Knowledge Discovery*, pp. 15–27, Springer, 2002.

[6] S. Ramaswamy, R. Rastogi, and K. Shim, "Efficient algorithms for mining outliers from large data sets," in *Proceedings of the 2000 ACM SIGMOD International Conference on Management of Data*, pp. 427–438, 2000.

[7] S. Wold, K. Esbensen, and P. Geladi, "Principal component analysis," *Chemometrics and Intelligent Laboratory Systems*, vol. 2, no. 1-3, pp. 37–52, 1987.

[8] H. Abdi and L. J. Williams, "Principal component analysis," *Wiley Interdisciplinary Reviews: Computational Statistics*, vol. 2, no. 4, pp. 433–459, 2010.

[9] G. A. Seber and A. J. Lee, *Linear regression analysis*, vol. 329. John Wiley & Sons, 2012.

[10] D. C. Montgomery, E. A. Peck, and G. G. Vining, *Introduction to linear regression analysis*. John Wiley & Sons, 2021.

[11] D. W. Hosmer Jr, S. Lemeshow, and R. X. Sturdivant, *Applied logistic regression*, vol. 398. John Wiley & Sons, 2013.

[12] S. Menard, *Applied logistic regression analysis*, vol. 106. Sage, 2002.

[13] I. Rish *et al.*, "An empirical study of the naive Bayes classifier," in *IJCAI 2001 Workshop on Empirical Methods in Artificial Intelligence*, vol. 3, pp. 41–46, 2001.

[14] M. M. Saritas and A. Yasar, "Performance analysis of ANN and Naive Bayes classification algorithm for data classification," *International Journal of Intelligent Systems and Applications in Engineering*, vol. 7, no. 2, pp. 88–91, 2019.

[15] Y. Freund and L. Mason, "The alternating decision tree learning algorithm," in *ICML*, vol. 99, pp. 124–133, Citeseer, 1999.

[16] J. Su and H. Zhang, "A fast decision tree learning algorithm," in *AAAI*, vol. 6, pp. 500–505, 2006.

[17] M. T. Hagan, H. B. Demuth, and M. Beale, *Neural network design*. PWS Publishing Co., 1997.

[18] S. I. Gallant and S. I. Gallant, *Neural network learning and expert systems*. MIT press, 1993.

[19] J. A. Bergou, U. Herzog, and M. Hillery, "11: Discrimination of quantum states," in *Quantum state estimation*, pp. 417–465, Springer, 2004.

[20] C.-W. Zhang, C.-F. Li, and G.-C. Guo, "General strategies for discrimination of quantum states," *Physics Letters A*, vol. 261, no. 1-2, pp. 25–29, 1999.

[21] J. A. Bergou, "Discrimination of quantum states," *Journal of Modern Optics*, vol. 57, no. 3, pp. 160–180, 2010.

[22] D. Qiu and L. Li, "Minimum-error discrimination of quantum states: Bounds and comparisons," *Physical Review A*, vol. 81, no. 4, p. 042329, 2010.

[23] S. M. Barnett and S. Croke, "Quantum state discrimination," *Advances in Optics and Photonics*, vol. 1, no. 2, pp. 238–278, 2009.

[24] J. Bae and L.-C. Kwek, "Quantum state discrimination and its applications," *Journal of Physics A: Mathematical and Theoretical*, vol. 48, no. 8, p. 083001, 2015.

[25] A. Chefles, "Quantum state discrimination," *Contemporary Physics*, vol. 41, no. 6, pp. 401–424, 2000.

[26] M. Piani and J. Watrous, "All entangled states are useful for channel discrimination," *Physical Review Letters*, vol. 102, no. 25, p. 250501, 2009.

[27] M. F. Sacchi, "Optimal discrimination of quantum operations," *Physical Review A*, vol. 71, no. 6, p. 062340, 2005.

[28] Q. Zhuang and S. Pirandola, "Ultimate limits for multiple quantum channel discrimination," *Physical Review Letters*, vol. 125, no. 8, p. 080505, 2020.

[29] H. Chernoff *et al.*, "A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations," *The Annals of Mathematical Statistics*, vol. 23, no. 4, pp. 493–507, 1952.

[30] W. Hoeffding, "Asymptotically optimal tests for multinomial distributions," *The Annals of Mathematical Statistics*, pp. 369–401, 1965.

[31] T. S. Han and K. Kobayashi, "The strong converse theorem for hypothesis testing," *IEEE Transactions on Information Theory*, vol. 35, no. 1, pp. 178–180, 1989.

[32] M. M. Wilde, M. Berta, C. Hirche, and E. Kaur, "Amortized channel divergence for asymptotic quantum channel discrimination," *Letters in Mathematical Physics*, vol. 110, no. 8, pp. 2277–2336, 2020.

[33] 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.

[34] A. Krawiec, Ł. Pawela, and Z. Puchała, "Discrimination of POVMs with rank-one effects," *Quantum Information Processing*, vol. 19, no. 12, pp. 1–12, 2020.

[35] Z. Puchała, Ł. Pawela, A. Krawiec, R. Kukulski, and M. Oszmaniec, "Multiple-shot and unambiguous discrimination of von Neumann measurements," *Quantum*, vol. 5, p. 425, 2021.

[36] 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, p. 3623, 2021.

[37] A. Krawiec, Ł. Pawela, and Z. Puchała, "Excluding false negative error in certification of quantum channels," *Scientific Reports*, vol. 11, no. 1, p. 21716, 2021.

[38] J. Watrous, *The Theory of Quantum Information*. Cambridge University Press, 2018.

[39] N. Brunner, M. Navascués, and T. Vértesi, "Dimension witnesses and quantum state discrimination," *Physical Review Letters*, vol. 110, no. 15, p. 150501, 2013.

[40] P. Horodecki, "Separability criterion and inseparable mixed states with positive partial transposition," *Physics Letters A*, vol. 232, no. 5, pp. 333–339, 1997.

[41] M. Enríquez, I. Wintrowicz, and K. Życzkowski, "Maximally entangled multipartite states: a brief survey," in *Journal of Physics: Conference Series*, vol. 698, p. 012003, IOP Publishing, 2016.

[42] D. Goyeneche, J. Bielawski, and K. Życzkowski, "Multipartite entanglement in heterogeneous systems," *Physical Review A*, vol. 94, no. 1, p. 012346, 2016.

[43] A. Burchardt, J. Czartowski, and K. Życzkowski, "Entanglement in highly symmetric multipartite quantum states," *Physical Review A*, vol. 104, no. 2, p. 022426, 2021.

[44] M. Kuś and K. Życzkowski, "Geometry of entangled states," *Physical Review A*, vol. 63, no. 3, p. 032307, 2001.

[45] I. Bengtsson and K. Życzkowski, *Geometry of quantum states: an introduction to quantum entanglement*. Cambridge University Press, 2017.

[46] M. Horodecki, "Entanglement measures.," *Quantum Inf. Comput.*, vol. 1, no. 1, pp. 3–26, 2001.

[47] M. Horodecki, P. Horodecki, and R. Horodecki, "Limits for entanglement measures," *Physical Review Letters*, vol. 84, no. 9, p. 2014, 2000.

[48] G. Vidal and R. F. Werner, "Computable measure of entanglement," *Physical Review A*, vol. 65, no. 3, p. 032314, 2002.

[49] V. Vedral, M. B. Plenio, M. A. Rippin, and P. L. Knight, "Quantifying entanglement," *Physical Review Letters*, vol. 78, no. 12, p. 2275, 1997.

[50] V. Vedral and M. B. Plenio, "Entanglement measures and purification procedures," *Physical Review A*, vol. 57, no. 3, p. 1619, 1998.

[51] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, "Quantum entanglement," *Reviews of Modern Physics*, vol. 81, no. 2, p. 865, 2009.

[52] K. Kraus, "General state changes in quantum theory," *Annals of Physics*, vol. 64, no. 2, pp. 311–335, 1971.

[53] M.-D. Choi, "Completely positive linear maps on complex matrices," *Linear Algebra and its Applications*, vol. 10, no. 3, pp. 285–290, 1975.

[54] E. Sudarshan, P. Mathews, and J. Rau, "Stochastic dynamics of quantum-mechanical systems," *Physical Review*, vol. 121, no. 3, p. 920, 1961.

[55] A. Jamiołkowski, "Linear transformations which preserve trace and positive semidefiniteness of operators," *Reports on Mathematical Physics*, vol. 3, no. 4, pp. 275–278, 1972.

[56] W. F. Stinespring, "Positive functions on C^*-algebras," *Proceedings of the American Mathematical Society*, vol. 6, no. 2, pp. 211–216, 1955.

[57] J. M. Renes, R. Blume-Kohout, A. J. Scott, and C. M. Caves, "Symmetric Informationally Complete Quantum Measurements," *Journal of Mathematical Physics*, vol. 45, no. 6, pp. 2171–2180, 2004.

[58] D. Appleby, "Symmetric Informationally Complete Measurements of Arbitrary Rank," *Optics and Spectroscopy*, vol. 103, pp. 416–428, 2007.

[59] G. S. Kopp, "SIC-POVMs and the Stark conjectures," *International Mathematics Research Notices*, vol. 2021, no. 18, pp. 13812–13838, 2021.

[60] I. J. Geng, K. Golubeva, and G. Gour, "What are the minimal conditions required to define a SIC POVM?," *arXiv preprint arXiv:2007.10483*, 2020.

[61] I. Bengtsson, "The number behind the simplest SIC POVM," *Foundations of Physics*, vol. 47, no. 8, pp. 1031–1041, 2017.

[62] P. Horodecki, Ł. Rudnicki, and K. Życzkowski, "Five open problems in quantum information theory," *PRX Quantum*, vol. 3, no. 1, p. 010101, 2022.

[63] M. Grassl, "On SIC-POVMs and MUBs in dimension 6," *arXiv preprint quant-ph/0406175*, 2004.

[64] S. T. Flammia, "On SIC-POVMs in prime dimensions," *Journal of Physics A: Mathematical and General*, vol. 39, no. 43, p. 13483, 2006.

[65] Z. Huangjun *et al.*, "SIC POVMs and Clifford groups in prime dimensions," *Journal of Physics. A, Mathematical and Theoretical (Online)*, vol. 43, 2010.

[66] C. W. Helstrom, *Quantum Detection and Estimation Theory*. Elsevier, 1976.

[67] G. Chiribella, G. M. D'Ariano, and P. Perinotti, "Theoretical framework for quantum networks," *Physical Review A*, vol. 80, no. 2, p. 022339, 2009.

[68] I. Nechita, Z. Puchała, Ł. Pawela, and K. Życzkowski, "Almost all quantum channels are equidistant," *Journal of Mathematical Physics*, vol. 59, no. 5, p. 052201, 2018.

[69] F. Hausdorff, "Der wertvorrat einer bilinearform," *Mathematische Zeitschrift*, vol. 3, no. 1, pp. 314–316, 1919.

[70] O. Toeplitz, "Das algebraische analogon zu einem satze von fejér," *Mathematische Zeitschrift*, vol. 2, no. 1, pp. 187–197, 1918.

[71] P. Gawron, Z. Puchała, J. A. Miszczak, Ł. Skowronek, and K. Życzkowski, "Restricted numerical range: a versatile tool in the theory of quantum information," *Journal of Mathematical Physics*, vol. 51, no. 10, p. 102204, 2010.

[72] Z. Puchała, P. Gawron, J. A. Miszczak, Ł. Skowronek, M.-D. Choi, and K. Życzkowski, "Product numerical range in a space with tensor product structure," *Linear Algebra and its Applications*, vol. 434, no. 1, pp. 327–342, 2011.

[73] E. Gutkin and K. Życzkowski, "Joint numerical ranges, quantum maps, and joint numerical shadows," *Linear Algebra and its Applications*, vol. 438, no. 5, pp. 2394–2404, 2013.

[74] C. F. Dunkl, P. Gawron, J. A. Holbrook, Z. Puchała, and K. Życzkowski, "Numerical shadows: measures and densities on the numerical range," *Linear Algebra and its Applications*, vol. 434, no. 9, pp. 2042–2080, 2011.

[75] C. F. Dunkl, P. Gawron, J. A. Holbrook, J. A. Miszczak, Z. Puchała, and K. Życzkowski, "Numerical shadow and geometry of quantum states," *Journal of Physics A: Mathematical and Theoretical*, vol. 44, no. 33, p. 335301, 2011.

[76] Z. Puchała, J. A. Miszczak, P. Gawron, C. F. Dunkl, J. A. Holbrook, and K. Życzkowski, "Restricted numerical shadow and the geometry of quantum entanglement," *Journal of Physics A: Mathematical and Theoretical*, vol. 45, no. 41, p. 415309, 2012.

[77] B. Collins, P. Gawron, A. E. Litvak, and K. Życzkowski, "Numerical range for random matrices," *Journal of Mathematical Analysis and Applications*, vol. 418, no. 1, pp. 516–533, 2014.

[78] N.-K. Tsing, "The constrained bilinear form and the C-numerical range," *Linear Algebra and its Applications*, vol. 56, pp. 195–206, 1984.

[79] C.-K. Li and H. Nakazato, "Some results on the q-numerical range," *Linear and Multilinear Algebra*, vol. 43, no. 4, pp. 385–409, 1998.

[80] C.-K. Li, "q-Numerical ranges of normal and convex matrices," *Linear and Multilinear Algebra*, vol. 43, no. 4, pp. 377–384, 1998.

[81] R. Duan, Y. Feng, and M. Ying, "Perfect distinguishability of quantum operations," *Physical Review Letters*, vol. 103, no. 21, p. 210501, 2009.

[82] G. Wang and M. Ying, "Unambiguous discrimination among quantum operations," *Physical Review A*, vol. 73, no. 4, p. 042301, 2006.

[83] W. Matthews, M. Piani, and J. Watrous, "Entanglement in channel discrimination with restricted measurements," *Physical Review A*, vol. 82, no. 3, p. 032302, 2010.

[84] M. Sedlák and M. Ziman, "Optimal single-shot strategies for discrimination of quantum measurements," *Physical Review A*, vol. 90, no. 5, p. 052312, 2014.

[85] M. Miková, M. Sedlák, I. Straka, M. Mičuda, M. Ziman, M. Ježek, M. Dušek, and J. Fiurášek, "Optimal entanglement-assisted discrimination of quantum measurements," *Physical Review A*, vol. 90, no. 2, p. 022317, 2014.

[86] J. Bae, D. Chruściński, and M. Piani, "More entanglement implies higher performance in channel discrimination tasks," *Physical Review Letters*, vol. 122, no. 14, p. 140404, 2019.

[87] 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.

[88] A. Jenčová and M. Plávala, "Conditions for optimal input states for discrimination of quantum channels," *Journal of Mathematical Physics*, vol. 57, no. 12, p. 122203, 2016.

[89] C. W. Helstrom, "Quantum detection and estimation theory," *Journal of Statistical Physics*, vol. 1, no. 2, pp. 231–252, 1969.

[90] M. Sedlák and M. Ziman, "Unambiguous comparison of unitary channels," *Physical Review A*, vol. 79, no. 1, p. 012303, 2009.

[91] A. Acín, "Statistical distinguishability between unitary operations," *Physical Review Letters*, vol. 87, no. 17, p. 177901, 2001.

[92] M. Ziman and M. Sedlák, "Single-shot discrimination of quantum unitary processes," *Journal of Modern Optics*, vol. 57, no. 3, pp. 253–259, 2010.

[93] R. Duan, Y. Feng, and M. Ying, "Local distinguishability of multipartite unitary operations," *Physical Review Letters*, vol. 100, no. 2, p. 020503, 2008.

[94] X.-F. Zhou, Y.-S. Zhang, and G.-C. Guo, "Unitary transformations can be distinguished locally," *Physical Review Letters*, vol. 99, no. 17, p. 170401, 2007.

[95] 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.

[96] L. Li and D. Qiu, "Local entanglement is not necessary for perfect discrimination between unitary operations acting on two qudits by local operations and classical communication," *Physical Review A*, vol. 77, no. 3, p. 032337, 2008.

[97] 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.

[98] K. M. Audenaert, J. Calsamiglia, R. Munoz-Tapia, E. Bagan, L. Masanes, A. Acin, and F. Verstraete, "Discriminating states: The quantum Chernoff bound," *Physical Review Letters*, vol. 98, no. 16, p. 160501, 2007.

[99] Z. Puchała, Ł. Pawela, and K. Życzkowski, "Distinguishability of generic quantum states," *Physical Review A*, vol. 93, no. 6, p. 062112, 2016.

[100] G. Chiribella, G. M. D'Ariano, and P. Perinotti, "Memory effects in quantum channel discrimination," *Physical Review Letters*, vol. 101, no. 18, p. 180501, 2008.

[101] 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.

[102] R. Duan, C. Guo, C.-K. Li, and Y. Li, "Parallel distinguishability of quantum operations," in *2016 IEEE International Symposium on Information Theory (ISIT)*, pp. 2259–2263, IEEE, 2016.

[103] 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.

[104] H. B. Dang, K. Blanchfield, I. Bengtsson, and D. M. Appleby, "Linear dependencies in Weyl–Heisenberg orbits," *Quantum Information Processing*, vol. 12, no. 11, pp. 3449–3475, 2013.

[105] M. Lundberg and L. Svensson, "The Haar measure and the generation of random unitary matrices," in *Processing Workshop Proceedings, 2004 Sensor Array and Multichannel Signal*, pp. 114–118, IEEE, 2004.

[106] N. J. Russell, L. Chakhmakhchyan, J. L. O'Brien, and A. Laing, "Direct dialling of Haar random unitary matrices," *New Journal of Physics*, vol. 19, no. 3, p. 033007, 2017.

[107] A. Chefles, "Unambiguous discrimination between linearly independent quantum states," *Physics Letters A*, vol. 239, no. 6, pp. 339–347, 1998.

[108] Y. Sun, J. A. Bergou, and M. Hillery, "Optimum unambiguous discrimination between subsets of nonorthogonal quantum states," *Physical Review A*, vol. 66, no. 3, p. 032315, 2002.

[109] Y. Feng, R. Duan, and M. Ying, "Unambiguous discrimination between mixed quantum states," *Physical Review A*, vol. 70, no. 1, p. 012308, 2004.

[110] U. Herzog and J. A. Bergou, "Optimum unambiguous discrimination of two mixed quantum states," *Physical Review A*, vol. 71, no. 5, p. 050301, 2005.

[111] M. Sedlák, M. Ziman, V. Bužek, and M. Hillery, "Unambiguous comparison of ensembles of quantum states," *Physical Review A*, vol. 77, no. 4, p. 042304, 2008.

[112] J. A. Bergou, U. Futschik, and E. Feldman, "Optimal unambiguous discrimination of pure quantum states," *Physical Review Letters*, vol. 108, no. 25, p. 250502, 2012.

[113] J. ur Rehman, A. Farooq, Y. Jeong, and H. Shin, "Quantum channel discrimination without entanglement," *Quantum Information Processing*, vol. 17, no. 10, p. 271, 2018.

[114] M. Ziman, T. Heinosaari, and M. Sedlák, "Unambiguous comparison of quantum measurements," *Physical Review A*, vol. 80, no. 5, p. 052102, 2009.

[115] G. Jaeger and A. Shimony, "Optimal distinction between two non-orthogonal quantum states," *Physics Letters A*, vol. 197, no. 2, pp. 83–87, 1995.

[116] J. Eisert, D. Hangleiter, N. Walk, I. Roth, D. Markham, R. Parekh, U. Chabaud, and E. Kashefi, "Quantum certification and benchmarking," *Nature Reviews Physics*, vol. 2, no. 7, pp. 382–390, 2020.

[117] M. Kliesch and I. Roth, "Theory of quantum system certification," *PRX quantum*, vol. 2, no. 1, p. 010201, 2021.

[118] C. Lu, J. Chen, and R. Duan, "Optimal perfect distinguishability between unitaries and quantum operations," *arXiv preprint arXiv:1010.2298*, 2010.

[119] C.-G. Ambrozie, "Finding positive matrices subject to linear restrictions," *Linear Algebra and its Applications*, vol. 426, no. 2-3, pp. 716–728, 2007.

[120] K. Fan, "Fixed-point and minimax theorems in locally convex topological linear spaces," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 38, no. 2, p. 121, 1952.

Appendix A

Proof of Theorem 2

In this appendix we will prove the Theorem 2. The proof requires a great deal of additional lemmas, therefore its scheme is presented in the following graph A.1.

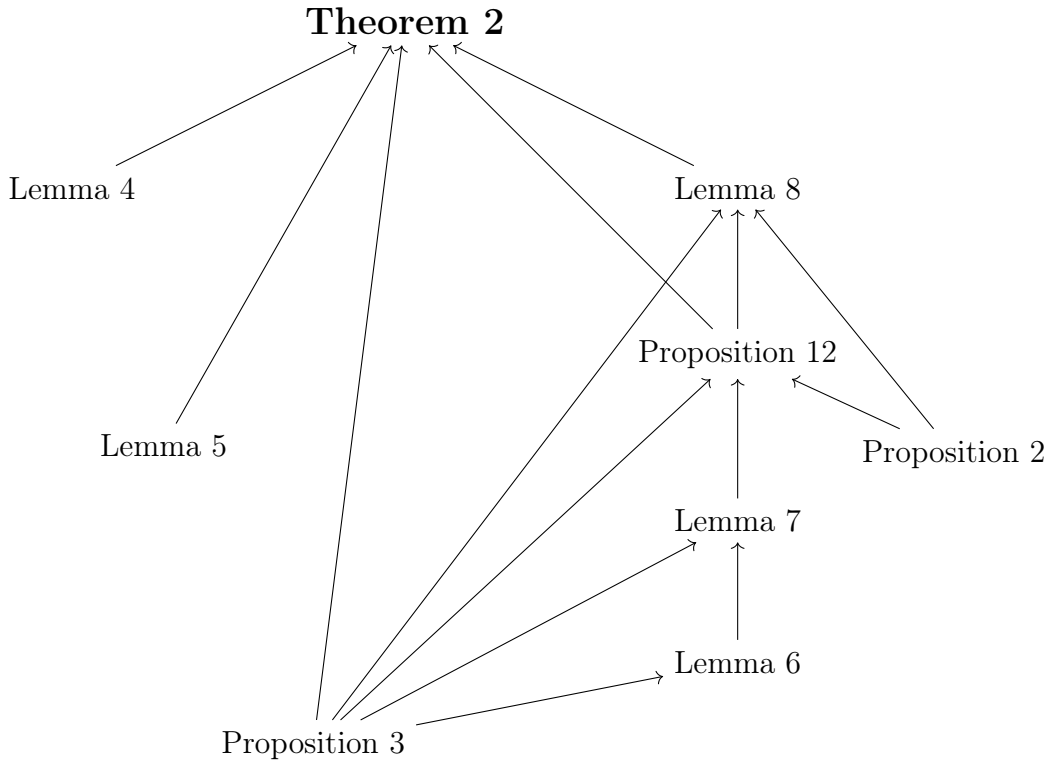

Figure A.1: Schematic representation of the proof of Theorem 2

Propositions 2 is a known result [38] and it is cited in Section 3.3. The proof

of Proposition 3, originally coming from [33], is presented in the main text in Section 3.3.

A.1 Proofs of technical lemmas and proposition

In this section we will begin with proving Lemmas 4 and 5. Next, we will prove Lemmas 6 and 7. We will continue with presenting the proofs of Proposition 12 and Lemma 8.

Lemma 4 *Let $U \in \mathcal{U}(\mathcal{X})$ and let $\mathcal{P}_U$ and $\mathcal{P}_{\mathbb{1}}$ be von Neumann measurements. Then, for every diagonal unitary matrix $E \in \mathcal{DU}(\mathcal{X})$ it holds that*

$$\|\mathcal{P}_U - \mathcal{P}_{\mathbb{1}}\|_\diamond \leq \|\Phi_{UE} - \Phi_{\mathbb{1}}\|_\diamond. \tag{A.1}$$

Proof. Let $\rho^\top \in \mathcal{D}(\mathcal{X})$ be a quantum state which satisfy the alternative formula for the diamond norm from Eq. (2.51), that is

$$\|\mathcal{P}_U - \mathcal{P}_{\mathbb{1}}\|_\diamond = \left\| \left(\mathbb{1} \otimes \sqrt{\rho^\top}\right) J\left(\mathcal{P}_U - \mathcal{P}_{\mathbb{1}}\right) \left(\mathbb{1} \otimes \sqrt{\rho^\top}\right) \right\|_1. \tag{A.2}$$

We calculate

$$\begin{aligned}
\|\mathcal{P}_U - \mathcal{P}_{\mathbb{1}}\|_\diamond &= \left\| \left(\mathbb{1} \otimes \sqrt{\rho^\top}\right) J\left(\mathcal{P}_U - \mathcal{P}_{\mathbb{1}}\right) \left(\mathbb{1} \otimes \sqrt{\rho^\top}\right) \right\|_1 \\
&= \left\| \left(\mathbb{1} \otimes \sqrt{\rho^\top}\right) \left(\sum_i |i\rangle\langle i| \otimes \left(|u_i\rangle\langle u_i| - |i\rangle\langle i|\right)^\top \right) \left(\mathbb{1} \otimes \sqrt{\rho^\top}\right) \right\|_1 \\
&= \left\| \sum_i |i\rangle\langle i| \otimes \sqrt{\rho^\top} \left(|u_i\rangle\langle u_i| - |i\rangle\langle i|\right)^\top \sqrt{\rho^\top} \right\|_1 \\
&= \left\| \sum_i |i\rangle\langle i| \otimes \sqrt{\rho^\top} \left(UE|i\rangle\langle i|E^\dagger U^\dagger - |i\rangle\langle i|\right)^\top \sqrt{\rho^\top} \right\|_1.
\end{aligned} \tag{A.3}$$

where the last equality follows from the simple observation that $|u_i\rangle\langle u_i| = UE|i\rangle\langle i|E^\dagger U^\dagger$.

Now, we will take advantage of the operational definition of the trace norm (see Eq. (2.24)) which yields

$$\|A\|_1 = \max_{V \in \mathcal{U}(\mathcal{X})} |\mathrm{tr} AV|. \tag{A.4}$$

Using this and the fact that the matrix is in a block-diagonal form we obtain

$$\begin{aligned}
&\left\| \sum_i |i\rangle\langle i| \otimes \sqrt{\rho^\top} \left(UE|i\rangle\langle i|E^\dagger U^\dagger - |i\rangle\langle i| \right)^\top \sqrt{\rho^\top} \right\|_1 \\
&= \sum_i \operatorname{tr} \left(\sqrt{\rho} \left(UE|i\rangle\langle i|E^\dagger U^\dagger - |i\rangle\langle i| \right) \sqrt{\rho} V_i \right) \\
&= \operatorname{tr} \left(\sum_i |i\rangle\langle i| \otimes \sqrt{\rho} \left(UE|i\rangle\langle i|E^\dagger U^\dagger - |i\rangle\langle i| \right) \sqrt{\rho} \right) \left(\sum_i |i\rangle\langle i| \otimes V_i \right),
\end{aligned} \tag{A.5}$$

where V_i is a unitary matrix that is optimal for i-th block. Therefore

$$\begin{aligned}
\|\mathcal{P}_U - \mathcal{P}_{\mathbb{1}}\|_\diamond &= \operatorname{tr} \left(\sum_i |i\rangle\langle i| \otimes \sqrt{\rho} \left(UE|i\rangle\langle i|E^\dagger U^\dagger - |i\rangle\langle i| \right) \sqrt{\rho} \right) \left(\sum_i |i\rangle\langle i| \otimes V_i \right) \\
&= \operatorname{tr} \left(\sum_{ij} |i\rangle\langle j| \otimes \sqrt{\rho} \left(UE|i\rangle\langle j|E^\dagger U^\dagger - |i\rangle\langle j| \right) \sqrt{\rho} \right) \left(\sum_i |i\rangle\langle i| \otimes V_i \right) \\
&\leq \max_{V \in \mathcal{U}(\mathcal{X} \otimes \mathcal{X})} \left| \operatorname{tr} \left(\sum_{ij} |i\rangle\langle j| \otimes \sqrt{\rho} \left(UE|i\rangle\langle j|E^\dagger U^\dagger - |i\rangle\langle j| \right) \sqrt{\rho} \right) V \right| \\
&= \left\| \sum_{ij} |i\rangle\langle j| \otimes \sqrt{\rho} \left(UE|i\rangle\langle j|E^\dagger U^\dagger - |i\rangle\langle j| \right) \sqrt{\rho} \right\|_1 \\
&= \left\| (\mathbb{1} \otimes \sqrt{\rho}) \left(\sum_{ij} |i\rangle\langle j| \otimes \left(UE|i\rangle\langle j|E^\dagger U^\dagger - |i\rangle\langle j| \right) \right) (\mathbb{1} \otimes \sqrt{\rho}) \right\|_1 .
\end{aligned} \tag{A.6}$$

Noting that

$$\begin{aligned}
&\sum_{ij} |i\rangle\langle j| \otimes \left(UE|i\rangle\langle j|E^\dagger U^\dagger - |i\rangle\langle j| \right) \\
&= \sum_{ij} |i\rangle\langle j| \otimes UE|i\rangle\langle j|E^\dagger U^\dagger - \sum_{ij} |i\rangle\langle j| \otimes |i\rangle\langle j| \\
&= (\mathbb{1} \otimes UE) \left(\sum_{ij} |i\rangle\langle j| \otimes |i\rangle\langle j| \right) \left(\mathbb{1} \otimes E^\dagger U^\dagger \right) - \sum_{ij} |i\rangle\langle j| \otimes |i\rangle\langle j| \\
&= (\mathbb{1} \otimes UE) \, |\mathbb{1}\rangle\rangle\langle\langle \mathbb{1}| \left(\mathbb{1} \otimes E^\dagger U^\dagger \right) - |\mathbb{1}\rangle\rangle\langle\langle \mathbb{1}| \\
&= |(UE)^\top\rangle\rangle\langle\langle (UE)^\top| - |\mathbb{1}\rangle\rangle\langle\langle \mathbb{1}|
\end{aligned} \tag{A.7}$$

we finally obtain

$$\begin{aligned}\|\mathcal{P}_U - \mathcal{P}_{\mathbf{1}}\|_\diamond &\leq \left\|(\mathbb{1}\otimes\sqrt{\rho})\left(|(UE)^\top\rangle\!\rangle\langle\!\langle(UE)^\top| - |\mathbb{1}\rangle\!\rangle\langle\!\langle\mathbb{1}|\right)(\mathbb{1}\otimes\sqrt{\rho})\right\|_1 \\ &\leq \|\Phi_{(UE)^\top} - \Phi_{\mathbf{1}}\|_\diamond = \|\Phi_{UE} - \Phi_{\mathbf{1}}\|_\diamond.\end{aligned} \tag{A.8}$$

■

Lemma 5 *Let $E_0 \in \mathcal{DU}(\mathcal{X})$, $U \in \mathcal{U}(\mathcal{X})$, $D(E) = \min_{\rho\in\mathcal{D}(\mathcal{X})} |\mathrm{Tr}\rho UE|$, $D(E_0) > 0$, λ_1, λ_d denote the eigenvalues of UE_0 such that the arc between them is the largest. Let P_1, P_d denote the projectors onto the subspaces spanned by the eigenvectors corresponding to λ_1, λ_d.*

Then, the function $|\mathrm{Tr}(\rho UE)|$ has a saddle point in (ρ_0, E_0) if and only if there exist states ρ_1, ρ_d such that

$$\rho_1 = P_1\rho_1P_1, \quad \rho_d = P_d\rho_dP_d, \quad \mathrm{diag}(\rho_1) = \mathrm{diag}(\rho_d). \tag{A.9}$$

Proof. We will begin with proving the reverse implication by defining a state $\rho_0 \in \mathcal{D}(\mathcal{X})$ as

$$\rho_0 \coloneqq \frac{1}{2}(\rho_1 + \rho_d) \tag{A.10}$$

We see that $|\mathrm{Tr}(UE_0\rho_0)| = D(E_0)$. For arbitrary $E \in \mathcal{DU}(\mathcal{X})$, direct calculation gives us

$$|\mathrm{Tr}(UE_0\rho_0)| \geq |\mathrm{Tr}(UE\rho_0)| \geq \min_{\rho\in\mathcal{D}(\mathcal{X})} |\mathrm{Tr}(UE\rho)|. \tag{A.11}$$

That means $D(E_0) \geq D(E)$ and $|\mathrm{Tr}(UE_0\rho_0)| = \min\limits_{\rho} |\mathrm{Tr}(UE_0\rho)| = \max\limits_{E} |\mathrm{Tr}(UE\rho_0)|$.

Now we prove the direct implication. Without loss of generality we may assume $\lambda_1 = \lambda$ and $\lambda_d = \overline{\lambda}$. Since ρ_0 gives minimum of the $|\mathrm{tr}\rho UE|$, thus ρ_0 is supported on the subspace spanned by the range of P_1 and P_d, i.e.

$$\rho_0 = P\rho_0P \text{ for } P = P_1 + P_d. \tag{A.12}$$

We may write

$$\rho_0 = P\rho_0P = P_1\rho_0P_1 + P_d\rho_0P_d + P_1\rho P_d + P_d\rho_0P_1 \tag{A.13}$$

and define

$$\begin{aligned}\rho_1 &= P_1\rho_0P_1, \\ \rho_d &= P_d\rho_0P_d, \\ \rho_{1d} &= P_1\rho_0P_d, \\ \rho_{d1} &= P_d\rho_0P_1.\end{aligned} \tag{A.14}$$

Note that the optimality forces $\mathrm{tr}(\rho_1) = \mathrm{tr}(\rho_d) = \frac{1}{2}$. Now we write

$$z_i = \langle i|\rho_0 U E_0|i\rangle = \lambda\langle i|\rho_1|i\rangle + \overline{\lambda}\langle i|\rho_d|i\rangle + 2\mathrm{Re}(\lambda\langle i|\rho_{d1}|i\rangle). \tag{A.15}$$

We have $\sum_i z_i = \frac{\lambda+\overline{\lambda}}{2}$. If elements z_i have different phases, then by additional diagonal unitary matrix one can increase the value of the sum and contradict to the fact that (ρ_0, E_0) is a saddle point. Therefore, we conclude that all elements have the same phase and therefore we obtain that $\langle i|\rho_1|i\rangle = \langle i|\rho_d|i\rangle$ for every i. ■

Lemma 6 *Von Neumann measurements $\mathcal{P}_U$ and $\mathcal{P}_{\mathbb{1}}$ can be discriminated perfectly if and only if for all real vectors $(x_1, \ldots, x_{2d}) \in \mathbb{R}^{2d}$ we have $0 \in W\left(\sum_{i=1}^{2d} x_i A_i\right)$.*

Proof. From Proposition 3 we know that $\mathcal{P}_U$ and $\mathcal{P}_{\mathbb{1}}$ can be discriminated perfectly if and only in there exists a state ρ satisfying $\mathrm{diag}(U^\dagger\rho) = 0$. We will write this condition as a semidefinite program, but first we need to introduce additional notation. Let $A_0 \coloneqq \mathbb{1}$ and

$$A_i \coloneqq U|i\rangle\langle i| + |i\rangle\langle i|U^\dagger \tag{A.16}$$

for $i = 1, \ldots, d$ and

$$A_i \coloneqq -\mathrm{i}\left(U|i\rangle\langle i| + |i\rangle\langle i|U^\dagger\right) \tag{A.17}$$

for $i = d+1, \ldots, 2d$. The semidefinite program for checking if von Neumann measurements can be discriminated perfectly is as follows

Primal problem		Dual problem	
maximize:	$\mathrm{Tr}\rho A_0$	minimize:	$\langle 0|Y|0\rangle$
subject to:	$\mathrm{Tr}\rho A_i = 0$	subject to:	$\sum_{i=0}^{2d} A_i Y_{ii} \geq \mathbb{1}$
	$\mathrm{Tr}\rho = 1$		$Y \in \mathrm{Herm}(\mathcal{X})$.
	$\rho \in \mathrm{Pos}(\mathcal{X})$		

In the above program, the target of maximization is a trivial function $\mathrm{tr}(\rho)$, which is later restricted to be equal to one. Thus, the primal problem reduces to satisfying the constraints. Fortunately, from [119, Thm 3] it holds that our primal problem has no solutions $\rho \geq 0$ if and only if

$$\inf_{(x_0,\ldots,x_{2d})\in\mathbb{R}^{2d+1}} \mathrm{e}^{x_0}\mathrm{tr}\left(\mathrm{e}^{\sum_{i=1}^{2d} x_i A_i}\right) - x_0 = -\infty. \tag{A.18}$$

This is equivalent to the condition that there exists a vector $(x_1, \ldots, x_{2d}) \in \mathbb{R}^{2d}$ such that $\sum_{i=1}^{2d} x_i A_i < 0$, that is, the real span of A_i contains only matrices without a determined sign. ■

Lemma 7 *Von Neumann measurements $\mathcal{P}_U$ and $\mathcal{P}_{\mathbb{1}}$ can be discriminated perfectly if and only if for every diagonal matrix D it holds that $0 \in W\left(UD + D^\dagger U^\dagger\right)$.*

Proof. In the first part of the proof we assume that $\mathcal{P}_U$ and $\mathcal{P}_{\mathbb{1}}$ can be discriminated perfectly. Therefore, from Proposition 3 we known that there must exist a state $\rho \in \mathcal{D}(\mathcal{X})$ such that

$$\operatorname{diag}(U^\dagger \rho) = 0. \tag{A.19}$$

Hence for every diagonal matrix D is must also hold that $\operatorname{diag}(D^\dagger U^\dagger \rho) = 0$ as well as $\operatorname{diag}(\rho U D) = 0$. Therefore $0 \in W\left(UD + D^\dagger U^\dagger\right)$.

To prove the other implication, we first note that every diagonal matrix of dimension $d = \dim(\mathcal{X})$ can be expressed as

$$D = \operatorname{diag}^\dagger(x_1 - \mathrm{i}x_{d+1}, x_2 - \mathrm{i}x_{d+2}, \ldots, x_d - \mathrm{i}x_{2d}) \tag{A.20}$$

for some real numbers $x_1, \ldots, x_{2d}$. We assume that for every diagonal matrix D it holds that $0 \in W\left(UD + D^\dagger U^\dagger\right)$. Thus, there must exist a state $|\psi\rangle$ such that

$$\langle\psi| \left(UD + D^\dagger U^\dagger\right) |\psi\rangle = 0. \tag{A.21}$$

Using the notation

$$A_i \coloneqq U|i\rangle\langle i| + |i\rangle\langle i|U^\dagger \tag{A.22}$$

for $i = 1, \ldots, d$ and

$$A_i \coloneqq -\mathrm{i}\left(U|i\rangle\langle i| + |i\rangle\langle i|U^\dagger\right) \tag{A.23}$$

for $i = d+1, \ldots, 2d$, the condition in Eq. (A.21) can be rewritten as

$$\langle\psi| \left(\sum_{i=1}^{2d} x_i A_i\right) |\psi\rangle = 0. \tag{A.24}$$

Using Lemma 6 we obtain that $\mathcal{P}_U$ and $\mathcal{P}_{\mathbb{1}}$ can be discriminated perfectly. ∎

Proposition 12 *Let $U \in \mathcal{U}(\mathcal{X})$. Von Neumann measurements $\mathcal{P}_U$ and $\mathcal{P}_{\mathbb{1}}$ can be discriminated perfectly if and only if for all diagonal unitary matrices $E \in \mathcal{DU}(\mathcal{X})$, unitary channels Φ_{UE} and $\Phi_{\mathbb{1}}$ can be discriminated perfectly.*

Proof. To prove the direct implication, assume that $\mathcal{P}_U$ and $\mathcal{P}_{\mathbb{1}}$ can be discriminated perfectly. Then, from Proposition 3 there must exist a quantum state $\rho \in \mathcal{D}(\mathcal{X})$ such that

$$\operatorname{diag}(U^\dagger \rho) = 0. \tag{A.25}$$

Therefore for every $E \in \mathcal{DU}(\mathcal{X})$ it also holds that $\operatorname{diag}(E^\dagger U^\dagger \rho) = 0$. Hence $0 \in W(E^\dagger U^\dagger)$, and by Proposition 2 we obtain that Φ_{UE} and $\Phi_{\mathbb{1}}$ can be discriminated perfectly

Now we will focus on the reverse implication. Assume that Φ_{UE} and the identity channel $\Phi_{\mathbf{1}}$ can be discriminated perfectly for every $E \in \mathcal{DU}(\mathcal{X})$. By Proposition 2 this can be written as $0 \in W(E^\dagger U^\dagger)$. We will be working towards showing that for any diagonal matrix D (not necessarily unitary), we have $0 \in W\left(UD + D^\dagger U^\dagger\right)$ (see Lemma 7). One may assume that D is invertible as otherwise we would have $\langle\varphi|\left(UD + D^\dagger U^\dagger\right)|\varphi\rangle = 0$ for some $|\varphi\rangle \in \ker(D)$. We can write

$$UD = UED_+, \tag{A.26}$$

where $E \in \mathcal{DU}(\mathcal{X})$ and D_+ is a strictly positive diagonal matrix. Let V be a unitary matrix such that

$$UE = V\mathrm{diag}^\dagger(\lambda)V^\dagger, \tag{A.27}$$

where λ is a vector of eigenvalues of UE. As we were assuming that $0 \in W(E^\dagger U^\dagger)$, there must exist a probability vector p, such that

$$\sum_i \lambda_i p_i = 0. \tag{A.28}$$

We define a quantum state σ as

$$\sigma = V\mathrm{diag}^\dagger(q)V^\dagger, \tag{A.29}$$

where

$$q_i = \frac{p_i}{\langle i|V^\dagger D_+ V|i\rangle}\left(\sum_j \frac{p_j}{\langle j|V^\dagger D_+ V|j\rangle}\right)^{-1}. \tag{A.30}$$

Using Eq. (A.26) and (A.27) we calculate

$$\begin{aligned}
\mathrm{tr}\left(UD\sigma\right) &= \mathrm{tr}\left(V\mathrm{diag}^\dagger(\lambda)V^\dagger D_+ V\mathrm{diag}^\dagger(q)V^\dagger\right) \\
&= \mathrm{tr}\left(V^\dagger D_+ V\mathrm{diag}^\dagger(q)V^\dagger V\mathrm{diag}^\dagger(\lambda)\right) \\
&= \sum_i \langle i|\left(V^\dagger D_+ V\mathrm{diag}^\dagger(q)\mathrm{diag}^\dagger(\lambda)\right)|i\rangle \\
&= \sum_i \lambda_i q_i \langle i|V^\dagger D_+ V|i\rangle,
\end{aligned} \tag{A.31}$$

and applying the definition of the state σ we have

$$\begin{aligned}\mathrm{tr}\,(UD\sigma) &= \sum_i \lambda_i \langle i|V^\dagger D_+ V|i\rangle \frac{p_i}{\langle i|V^\dagger D_+ V|i\rangle}\left(\sum_j \frac{p_j}{\langle j|V^\dagger D_+ V|j\rangle}\right)^{-1} \\ &= \sum_i \lambda_i p_i \left(\sum_j \frac{p_j}{\langle j|V^\dagger D_+ V|j\rangle}\right)^{-1} = 0,\end{aligned} \tag{A.32}$$

where the last equality follows from Eq. (A.28). Therefore, $0 \in W\,(UD)$, and eventually $0 \in W\left(UD + D^\dagger U^\dagger\right)$. ∎

Lemma 8 *Let $U \in \mathcal{U}(\mathcal{X})$. Then*

$$\min_{\rho\in\mathcal{D}(\mathcal{X})} \max_{E\in\mathcal{DU}(\mathcal{X})} |\mathrm{Tr}(\rho UE)| = \max_{E\in\mathcal{DU}(\mathcal{X})} \min_{\rho\in\mathcal{D}(\mathcal{X})} |\mathrm{Tr}(\rho UE)|. \tag{A.33}$$

Proof. By the use of Proposition 2 we note that

$$\min_{E\in\mathcal{DU}(\mathcal{X})} \|\Phi_{UE} - \Phi_{\mathbf{1}}\|_\diamond = 2\sqrt{1 - \max_{E\in\mathcal{DU}(\mathcal{X})} \min_{\rho\in\mathcal{D}(\mathcal{X})} |\mathrm{Tr}\,(\rho UE)\,|^2}. \tag{A.34}$$

Let us first focus on the case when $\min_{E\in\mathcal{DU}(\mathcal{X})} \|\Phi_{UE} - \Phi_{\mathbf{1}}\|_\diamond = 2$. Then

$$\max_{E\in\mathcal{DU}(\mathcal{X})} \min_{\rho\in\mathcal{D}(\mathcal{X})} |\mathrm{Tr}\,(\rho UE)\,| = 0. \tag{A.35}$$

On the other hand, from Proposition 12 we have that $\mathcal{P}_U$ and $\mathcal{P}_1$ can be discriminated perfectly. Therefore, following Proposition 3, there exists a state $\rho_0 \in \mathcal{D}(\mathcal{X})$ such that $\mathrm{diag}\,(\rho_0 U) = 0$, and hence $|\mathrm{Tr}(\rho_0 UE)| = 0$ for every $E \in \mathcal{DU}(\mathcal{X})$. Thus

$$0 = \max_{E\in\mathcal{DU}(\mathcal{X})} |\mathrm{Tr}(\rho_0 UE)| = \min_{\rho\in\mathcal{D}(\mathcal{X})} \max_{E\in\mathcal{DU}(\mathcal{X})} |\mathrm{Tr}(\rho UE)|. \tag{A.36}$$

Therefore, in the case when $\min_{E\in\mathcal{DU}(\mathcal{X})} \|\Phi_{UE} - \Phi_{\mathbf{1}}\|_\diamond = 2$, we have

$$\max_{E\in\mathcal{DU}(\mathcal{X})} \min_{\rho\in\mathcal{D}(\mathcal{X})} |\mathrm{Tr}\,(\rho UE)\,| = \min_{\rho\in\mathcal{D}(\mathcal{X})} \max_{E\in\mathcal{DU}(\mathcal{X})} |\mathrm{Tr}(\rho UE)|. \tag{A.37}$$

Now we assume assume that $\min_{E\in\mathcal{DU}(\mathcal{X})} \|\Phi_{UE} - \Phi_{\mathbf{1}}\|_\diamond < 2$. From Proposition 2 we have

$$\min_{E\in\mathcal{DU}(\mathcal{X})} \|\Phi_{UE} - \Phi_{\mathbf{1}}\|_\diamond = 2\sqrt{1 - \max_{E\in\mathcal{DU}(\mathcal{X})} \min_{\rho\in\mathcal{D}(\mathcal{X})} |\mathrm{Tr}(\rho UE)|^2}. \tag{A.38}$$

In the case of $\rho_0 \in \mathcal{D}(\mathcal{X})$ and $E_0 \in \mathcal{DU}(\mathcal{X})$ which saturate $\min_{E\in\mathcal{DU}(\mathcal{X})} \|\Phi_{UE} - \Phi_{\mathbf{1}}\|_\diamond$, we have that $0 \notin W(UE_0)$.

Let $\mathcal{D}^{\leq 1}(\mathcal{X})$ be the set of diagonal matrices E such that $|E_{ii}| \leq 1$ for every i. Recall from Subsection 2.1.1 that the set $\mathcal{D}(\mathcal{X})$ is compact and convex. Similarly, the set $\mathcal{D}^{\leq 1}(\mathcal{X})$ is also convex and compact. Moreover, the sets $\{E \in \mathcal{D}^{\leq 1}(\mathcal{X}) : \mathrm{Re}(\mathrm{Tr}(\rho UE)) = \max_{D\in\mathcal{D}^{\leq 1}(\mathcal{X})} \mathrm{Re}(\mathrm{Tr}(\rho UD))\}$ and $\{\rho \in \mathcal{D}(\mathcal{X}) : \mathrm{Re}(\mathrm{Tr}(\rho UE)) = \min_{\sigma\in\mathcal{D}(\mathcal{X})} \mathrm{Re}(\mathrm{Tr}(\sigma UE))\}$ are convex. Next, we check that all the assumptions of the Theorem 3 in [120] are fulfilled, thus we obtain the existence of saddle points, that is

$$\min_{\rho\in\mathcal{D}(\mathcal{X})} \max_{E\in\mathcal{D}^{\leq 1}(\mathcal{X})} \mathrm{Re}\left(\mathrm{Tr}(\rho UE)\right) = \max_{E\in\mathcal{D}^{\leq 1}(\mathcal{X})} \min_{\rho\in\mathcal{D}(\mathcal{X})} \mathrm{Re}\left(\mathrm{Tr}(\rho UE)\right). \tag{A.39}$$

From the above equation, for a saddle point (ρ_0, E_0) we have

$$\mathrm{Re}\left(\mathrm{Tr}\left(\rho_0 UE_0\right)\right) = \mathrm{Tr}\left(\rho_0 UE_0\right) = \left|\mathrm{Tr}\left(\rho_0 UE_0\right)\right|. \tag{A.40}$$

Moreover,

$$\max_E \left|\mathrm{Tr}\left(\rho_0 UE\right)\right| = \sum_i \left|\langle i|\rho_0 U|i\rangle\right| = \mathrm{Tr}\left(\rho_0 UE_0\right) \tag{A.41}$$

and

$$\mathrm{Tr}\left(\rho_0 UE_0\right) = \min_\rho \left|\mathrm{Tr}\left(\rho UE_0\right)\right|. \tag{A.42}$$

From the above we obtain that (ρ_0, E_0) is the saddle point of $|\mathrm{Tr}(\rho UE)|$, that is

$$\min_{\rho\in\mathcal{D}(\mathcal{X})} \max_{E\in\mathcal{D}^{\leq 1}(\mathcal{X})} |\mathrm{Tr}(\rho UE)| = \max_{E\in\mathcal{D}^{\leq 1}(\mathcal{X})} \min_{\rho\in\mathcal{D}(\mathcal{X})} |\mathrm{Tr}(\rho UE)|. \tag{A.43}$$

Let $E_0 \coloneqq F_0 D$, where $F_0 \in \mathcal{DU}(\mathcal{X})$ and D is a diagonal matrix with $0 \leq D_{ii} \leq 1$ for every i. We will show that we have the saddle point also for (ρ_0, F_0). First, observe that for every $U \in \mathcal{U}(\mathcal{X})$ it hods that

$$\min_\rho |\mathrm{Tr}\rho U| \geq \min_\rho |\mathrm{Tr}\rho UD|. \tag{A.44}$$

Now, we will consider two cases: when $0 \in W(U)$ and when $0 \notin W(U)$. In the former case, when $0 \in W(U)$, then for some probability vector p we have $\sum_i \lambda_i p_i = 0$, where λ_i are the eigenvalues of U. If there exists i such that $\langle\lambda_i|D|\lambda_i\rangle = 0$, then $|\mathrm{Tr}|\lambda_i\rangle\langle\lambda_i|UD| = 0$. Otherwise, we can take the state $\rho = \sum_i q_i |\lambda_i\rangle\langle\lambda_i|$, where $q_i = \frac{p_i}{\langle\lambda_i|D|\lambda_i\rangle}$ and notice that $0 \in W(UD)$.

In the case when $0 \notin W(U)$, for the most distant pair of eigenvalues, λ_1, λ_d, of the matrix U we can the the Töplitz-Hausdorff theorem. It gives the inclusion of

the interval in a numerical range

$$[\mathrm{Tr}|\lambda_1\rangle\langle\lambda_1|UD, \mathrm{Tr}|\lambda_d\rangle\langle\lambda_d|UD] = [\lambda_1\langle\lambda_1|D|\lambda_1\rangle, \lambda_d\langle\lambda_d|D|\lambda_d\rangle] \subset W(UD). \quad \text{(A.45)}$$

In our case, using the optimality condition we obtain

$$\min_{\rho} |\mathrm{Tr}\rho UF_0| = \min_{\rho} |\mathrm{Tr}\rho UF_0D|. \quad \text{(A.46)}$$

To finish the proof, it remains to check whether (ρ_0, F_0) is the saddle point. We have

$$\begin{aligned} |\mathrm{Tr}\rho_0 UF_0| &\leq \max_{E\in\mathcal{D}^{\leq 1}(\mathcal{X})} |\mathrm{Tr}\rho_0 UE| = |\mathrm{Tr}\rho_0 UE_0| = \min_{\rho} |\mathrm{Tr}\rho UF_0D| \\ &= \min_{\rho} |\mathrm{Tr}\rho UF_0| \leq |\mathrm{Tr}\rho_0 UF_0|. \end{aligned} \quad \text{(A.47)}$$

From the above we obtain

$$|\mathrm{Tr}\rho_0 UF_0| = \min_{\rho} |\mathrm{Tr}\rho UF_0| = \max_{E\in\mathcal{D}^{\leq 1}(\mathcal{X})} |\mathrm{Tr}\rho_0 UE|, \quad \text{(A.48)}$$

and eventually

$$\min_{\rho\in\mathcal{D}(\mathcal{X})} \max_{E\in\mathcal{DU}(\mathcal{X})} |\mathrm{Tr}(\rho UE)| = \max_{E\in\mathcal{DU}(\mathcal{X})} \min_{\rho\in\mathcal{D}(\mathcal{X})} |\mathrm{Tr}(\rho UE)|. \quad \text{(A.49)}$$

■

A.2 Proof of Theorem 2

Proof of Theorem 2. We will consider two cases, when $\min_{E\in\mathcal{DU}(\mathcal{X})} \|\Phi_{UE} - \Phi_{\mathbf{1}}\|_\diamond = 2$, and when $\min_{E\in\mathcal{DU}(\mathcal{X})} \|\Phi_{UE} - \Phi_{\mathbf{1}}\|_\diamond < 2$. The former case is considered in Proposition 12, which states that measurements $\mathcal{P}_U$ and $\mathcal{P}_{\mathbf{1}}$ can be discriminated perfectly if and only if unitary channels Φ_{UE} and $\Phi_{\mathbf{1}}$ can be discriminated perfectly for every $E \in \mathcal{DU}(\mathcal{X})$. Therefore, when $\min_{E\in\mathcal{DU}(\mathcal{X})} \|\Phi_{UE} - \Phi_{\mathbf{1}}\|_\diamond = 2$, then also $\|\mathcal{P}_U - \mathcal{P}_{\mathbf{1}}\|_\diamond = 2$.

In the remaining part of the proof we will focus on the latter case when $\min_{\mathcal{DU}(\mathcal{X})} \|\Phi_{UE} - \Phi_{\mathbf{1}}\|_\diamond < 2$. Lemma 8 guarantees the existence of a saddle point (ρ_0, E_0) and utilizing Lemma 5 (and its proof), we define a new state

$$\tau = \frac{1}{2}(\rho_1 + \rho_d). \quad \text{(A.50)}$$

Following the proof of Proposition 3 we calculate according to Eq. (3.11).

$$\left\|(\mathbb{1}\otimes\sqrt{\tau})J\left(\mathcal{P}_{UE_0}-\mathcal{P}_{\mathbf{1}}\right)(\mathbb{1}\otimes\sqrt{\tau})\right\|_1=\sum_{i=1}^{d}\sqrt{\left(\langle u_i|\tau|u_i\rangle+\langle i|\tau|i\rangle\right)^2-4\left|\langle u_i|\tau|i\rangle\right|^2}. \tag{A.51}$$

Direct calculation gives

$$\sum_{i=1}^{d}\sqrt{\left(\langle u_i|\tau|u_i\rangle+\langle i|\tau|i\rangle\right)^2-4\left|\langle u_i|\tau|i\rangle\right|^2}=2\sqrt{1-\left|\frac{\lambda_1+\lambda_d}{2}\right|^2}, \tag{A.52}$$

where $\left|\frac{\lambda_1+\lambda_d}{2}\right|=\left|\mathrm{Tr}\left(\tau U E_0\right)\right|$. To finish this proof we use Lemma 4, which states that $\|\mathcal{P}_U-\mathcal{P}_{\mathbf{1}}\|_\diamond\leq\|\Phi_{UE}-\Phi_{\mathbf{1}}\|_\diamond$. Therefore

$$\begin{aligned}2\sqrt{1-\left|\frac{\lambda_1+\lambda_d}{2}\right|^2}&=\left\|(\mathbb{1}\otimes\sqrt{\tau})J(\mathcal{P}_U-\mathcal{P}_{\mathbf{1}})(\mathbb{1}\otimes\sqrt{\tau})\right\|_1\leq\|\mathcal{P}_U-\mathcal{P}_{\mathbf{1}}\|_\diamond\\&\leq\min_{E\in\mathcal{DU}(\mathcal{X})}\|\Phi_{UE}-\Phi_{\mathbf{1}}\|_\diamond=2\sqrt{1-\left|\frac{\lambda_1+\lambda_d}{2}\right|^2}\end{aligned} \tag{A.53}$$

which eventually gives $\|\mathcal{P}_U-\mathcal{P}_{\mathbf{1}}\|_\diamond=\min_{E\in\mathcal{DU}(\mathcal{X})}\|\Phi_{UE}-\Phi_{\mathbf{1}}\|_\diamond$. ■

Appendix B

Proof of Theorem 11

In this appendix we will present the proof of Theorem 11. Before we do so, we will quickly review the results for asymmetric discrimination of pure quantum states and unitary channels in Section B.1 These results will be used to prove the Theorem 11, and we will state them along with a short background of the problem formulation.

Later, in Section B.2, we will prove some additional lemmas which will make the proof of Theorem 11 easier to read. Finally, the proof of Theorem 11 will be presented in Section B.3.

B.1 Asymmetric discrimination of pure states and unitary channels

Asymmetric discrimination of pure states [36,66] Consider the problem of asymmetric discrimination of pure quantum states $|\psi\rangle$ and $|\varphi\rangle$. Let $|\psi\rangle$ correspond to the H_0 hypothesis while $|\varphi\rangle$ correspond to the H_1 hypothesis. The scheme of discrimination of pure states is straightforward - we only measure the state with the binary measurement $\{\Omega, \mathbb{1} - \Omega\}$. When the obtained measurement label corresponds to the effect Ω, we accept the hypothesis H_0. When the measurement label corresponds to the effect $\mathbb{1} - \Omega$, we reject the H_0 hypothesis.

The probability of making the false positive error yields

$$\alpha(\Omega) \coloneqq \langle\psi| (\mathbb{1} - \Omega) |\psi\rangle \tag{B.1}$$

The optimal probability of making the false negative error is defined as

$$\beta_\delta \coloneqq \min_{\Omega} \{\langle\varphi|\Omega|\varphi\rangle : \alpha(\Omega) \leq \delta\}. \tag{B.2}$$

Proposition 13 *With the notation as above, it holds that*

$$\beta_\delta = \begin{cases} 0, & \text{if } |\langle\psi|\varphi\rangle| \leq \sqrt{\delta} \\ \left(|\langle\psi|\varphi\rangle|\sqrt{1-\delta} - \sqrt{1-|\langle\psi|\varphi\rangle|^2}\sqrt{\delta}\right)^2, & \text{if } |\langle\psi|\varphi\rangle| > \sqrt{\delta} \end{cases} . \tag{B.3}$$

Asymmetric discrimination of unitary channels [36, 118] Let H_0 hypothesis correspond to the identity channel $\Phi_{\mathbb{1}} = \mathbb{1}$ and let H_0 correspond to another unitary channel Φ_U. To discriminate unitary channels, we can use an additional register and prepare an entangled input state. We apply the the unitary channel (either $\Phi_{\mathbb{1}}$ or Φ_U) on the first register and the identity channels on the second register. Having the input state $|\psi\rangle$ fixed, we can write conditional hypotheses as $H_0 : |\psi\rangle$ and $H_1 : (U \otimes \mathbb{1})|\psi\rangle$.

Then, we measure the output state with a binary measurement $\mathcal{P}_F = \{\Omega, \mathbb{1} - \Omega\}$. When we obtained a measurement label corresponding to the effect Ω, then we accept the hypothesis H_0. When we obtained the measurement label corresponding to the effect $\mathbb{1} - \Omega$, then we reject the H_0 hypothesis and accept the H_1 hypothesis.

For the fixed input state and final measurement, the probability of making the false positive error yields

$$\alpha(\psi, \Omega) \coloneqq \langle\psi| (\mathbb{1} - \Omega) |\psi\rangle. \tag{B.4}$$

The optimized probability of making the false negative error yields

$$\beta_\delta \coloneqq \min_{\psi,\Omega} \left\{ \operatorname{tr} \left(\Omega \left(\Phi_U \otimes \mathbb{1}\right) \left(|\psi\rangle\langle\psi|\right)\right) : \ \alpha(\psi, \Omega) \leq \delta \right\}. \tag{B.5}$$

Proposition 14 *With the notation as above, it holds that*

$$\beta_\delta = \nu^2_{\sqrt{1-\delta}}(U). \tag{B.6}$$

B.2 Proofs of technical lemmas

Both lemmas presented in this section were proved in [36].

Lemma 9 *Let $\delta > 0$ and Ω be a measurement effect, that is a positive semidefinite operator satisfying $\Omega \leq \mathbb{1}$. For every quantum channel Ψ and quantum states ρ_0, ρ_1 it holds that*

$$\min_{\Omega:\ \operatorname{tr}(\Omega\rho_0) \geq 1-\delta} \operatorname{tr} \left(\Omega\rho_1\right) \leq \min_{\Omega:\ \operatorname{tr}(\Omega\Psi(\rho_0)) \geq 1-\delta} \operatorname{tr} \left(\Omega\Psi(\rho_1)\right). \tag{B.7}$$

Proof. First we note that

$$\min_{\Omega:\ \mathrm{tr}(\Omega\Psi(\rho_0))\geq 1-\delta} \mathrm{tr}\left(\Omega\Psi(\rho_1)\right) = \min_{\Omega:\ \mathrm{tr}\left(\Psi^\dagger(\Omega)\rho_0\right)\geq 1-\delta} \mathrm{tr}\left(\Psi^\dagger(\Omega)\rho_0\right) \tag{B.8}$$

and $\Psi^\dagger(\Omega)$ is also a measurement effect. Moreover

$$\left\{\Psi^\dagger(\Omega):\ \mathrm{tr}\left(\Psi^\dagger(\Omega)\rho_0\right)\geq 1-\delta\right\} \subseteq \left\{\Omega:\ \mathrm{tr}\left(\Omega\rho_0\right)\geq 1-\delta\right\}. \tag{B.9}$$

Finally

$$\min_{\Omega:\ \mathrm{tr}(\Omega\rho_0)\geq 1-\delta} \mathrm{tr}\left(\Omega\rho_1\right) \leq \min_{\Omega:\ \mathrm{tr}(\Omega\Psi(\rho_0))\geq 1-\delta} \mathrm{tr}\left(\Omega\Psi(\rho_1)\right). \tag{B.10}$$

■

Lemma 10 *Let $\rho_0 = \frac{1}{2}\rho_1 + \frac{1}{2}\rho_d$ a quantum state satisfying conditions given by Lemma 5 in Appendix A. Then, for each $i \in \{1,\ldots,d\}$ it holds that*

$$\mathrm{tr}\left(\sqrt{\rho_0}|i\rangle\langle i|\sqrt{\rho_0}\right) = \mathrm{tr}\left(\sqrt{\rho_0}U|i\rangle\langle i|U^\dagger\sqrt{\rho_0}\right). \tag{B.11}$$

Moreover, for each $i \in \{1,\ldots,d\}$ such that $\langle i|\rho_0|i\rangle \neq 0$ it holds that

$$\left|\frac{\langle i|\rho_0 U|i\rangle}{\langle i|\rho_0|i\rangle}\right| = \left|\frac{\lambda_1+\lambda_d}{2}\right|. \tag{B.12}$$

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

$$\begin{aligned}
\mathrm{tr}\left(\sqrt{\rho_0}U|i\rangle\langle i|U^\dagger\sqrt{\rho_0}\right) &= \langle i|U^\dagger\rho U|i\rangle = \langle i|U^\dagger\left(\frac{1}{2}\rho_1+\frac{1}{2}\rho_d\right)U|i\rangle \\
&= \langle i|U^\dagger\left(\frac{1}{2}\Pi_1\rho_1\Pi_1+\frac{1}{2}\Pi_d\rho_d\Pi_d\right)U|i\rangle \\
&= \langle i|\left(\sum_{i=1}^d \overline{\lambda_i}\Pi_i^\dagger\right)\left(\frac{1}{2}\Pi_1\rho_1\Pi_1+\frac{1}{2}\Pi_d\rho_d\Pi_d\right)\left(\sum_{i=1}^d \lambda_i\Pi_i\right)|i\rangle \\
&= \langle i|\left(\frac{1}{2}\rho_1+\frac{1}{2}\rho_d\right)|i\rangle = \mathrm{tr}\left(\sqrt{\rho_0}|i\rangle\langle i|\sqrt{\rho_0}\right),
\end{aligned} \tag{B.13}$$

where the third equality follows from Lemma 5 in Appendix A.

To prove the second part of the proposition we calculate

$$\begin{aligned}\left|\frac{\langle i|\rho_0 U|i\rangle}{\langle i|\rho_0|i\rangle}\right| &= \left|\frac{\langle i|\left(\frac{1}{2}\rho_1+\frac{1}{2}\rho_d\right)\left(\sum_{i=1}^{d}\lambda_i\Pi_i\right)|i\rangle}{\langle i|\rho_0|i\rangle}\right| \\ &= \left|\frac{\langle i|\sum_{i=1}^{d}\lambda_i\left(\frac{1}{2}\Pi_1\rho_1\Pi_1+\frac{1}{2}\Pi_d\rho_d\Pi_d\right)\Pi_i|i\rangle}{\langle i|\rho_0|i\rangle}\right| \\ &= \left|\frac{\langle i|\left(\frac{1}{2}\lambda_1\Pi_1\rho_1\Pi_1+\frac{1}{2}\lambda_d\Pi_d\rho_d\Pi_d\right)|i\rangle}{\langle i|\rho_0|i\rangle}\right| \\ &= \left|\frac{\langle i|\left(\frac{1}{2}\lambda_1\rho_1+\frac{1}{2}\lambda_d\rho_d\right)|i\rangle}{\langle i|\rho_0|i\rangle}\right| = \left|\frac{\lambda_1+\lambda_d}{2}\right|.\end{aligned} \tag{B.14}$$

■

B.3 Proof of Theorem 11

The following proof was originally presented in [36].

Proof. In the scheme of certification of von Neumann measurements, the optimized probability of type II error can be expressed as

$$\beta_\delta \coloneqq \min_{\psi,\Omega}\left\{\mathrm{tr}\left(\Omega\left(\mathcal{P}_U\otimes\mathbb{1}\right)\left(|\psi\rangle\langle\psi|\right)\right) : \alpha(\psi,\Omega)\leq\delta\right\}. \tag{B.15}$$

Our goal is to prove that

$$\beta_\delta = \max_{E\in\mathcal{DU}(\mathcal{X})}\nu^2_{\sqrt{1-\delta}}\left(UE\right). \tag{B.16}$$

The proof is divided into two parts. In the first part we will show the lower bound on β_δ using Data Processing Inequality presented in Lemma 9. In the second part we will show the upper bound on β_δ.

The lower bound

This part of the proof will mostly be based on Data Processing Inequality in Lemma 9. To show that

$$\beta_\delta \geq \max_{E\in\mathcal{DU}(\mathcal{X})}\nu^2_{\sqrt{1-\delta}}\left(UE\right), \tag{B.17}$$

let us begin with an observation that every von Neumann measurement $\mathcal{P}_U$ can be rewritten as $\Delta \circ \Phi_{(UE)^\dagger}$, where Δ denotes the completely dephasing channel and $E \in \mathcal{DU}(\mathcal{X})$. Therefore, utilizing the Data Processing Inequality in Lemma 9, along with the certification scheme of unitary channels in Proposition 14, the optimized probability of the type II error is lower-bounded by

$$\begin{aligned}\beta_\delta &\geq \min_{\psi,\Omega} \left\{ \mathrm{tr}\left(\Omega \left(\Phi_{(UE)^\dagger} \otimes \mathbb{1}\right)(|\psi\rangle\langle\psi|)\right) : \alpha(\psi,\Omega) \leq \delta \right\} \\ &= \nu^2_{\sqrt{1-\delta}}\left((UE)^\dagger\right) = \nu^2_{\sqrt{1-\delta}}\left(UE\right),\end{aligned} \tag{B.18}$$

which holds for each $E \in \mathcal{DU}(\mathcal{X})$. Hence, maximizing the value of $\nu^2_{\sqrt{1-\delta}}(UE)$ over $E \in \mathcal{DU}(\mathcal{X})$ leads to the lower bound of the form

$$\beta_\delta \geq \max_{E \in \mathcal{DU}(\mathcal{X})} \nu^2_{\sqrt{1-\delta}}\left(UE\right). \tag{B.19}$$

The upper bound

Now we proceed to proving the upper bound.The proof of the inequality

$$\beta_\delta \leq \max_{E \in \mathcal{DU}(\mathcal{X})} \nu^2_{\sqrt{1-\delta}}\left(UE\right) \tag{B.20}$$

will be divided into two cases depending on the diamond norm distance between $\mathcal{P}_U$ and $\mathcal{P}_{\mathbb{1}}$. In either case we will construct a strategy, that is choose a state $|\psi_0\rangle$ and a measurement Ω_0. As for every choice of $|\psi\rangle$ and Ω it holds that

$$\beta_\delta \leq \mathrm{tr}\left(\Omega(\mathcal{P}_U \otimes \mathbb{1})(|\psi\rangle\langle\psi|)\right), \tag{B.21}$$

we will show that for some fixed $|\psi_0\rangle$ and Ω_0 it holds that

$$\mathrm{tr}\left(\Omega_0(\mathcal{P}_U \otimes \mathbb{1})(|\psi_0\rangle\langle\psi_0|)\right) = \max_{E \in \mathcal{DU}(\mathcal{X})} \nu^2_{\sqrt{1-\delta}}\left(UE\right). \tag{B.22}$$

First we focus on the case when $\|\mathcal{P}_U - \mathcal{P}_{\mathbb{1}}\|_\diamond = 2$. We take a state $|\psi_0\rangle$ for which it holds that

$$\|\mathcal{P}_U - \mathcal{P}_{\mathbb{1}}\|_\diamond = \|\left((\mathcal{P}_U - \mathcal{P}_{\mathbb{1}}) \otimes \mathbb{1}\right)(|\psi_0\rangle\langle\psi_0|)\|_1. \tag{B.23}$$

Then, the output states $(\mathcal{P}_U \otimes \mathbb{1})(|\psi_0\rangle\langle\psi_0|)$ and $(\mathcal{P}_{\mathbb{1}} \otimes \mathbb{1})(|\psi_0\rangle\langle\psi_0|)$ are orthogonal and by taking the measurement Ω_0 as the projection onto the support of $(\mathcal{P}_{\mathbb{1}} \otimes \mathbb{1})(|\psi_0\rangle\langle\psi_0|)$ we obtain

$$\mathrm{tr}\left(\Omega_0(\mathcal{P}_U \otimes \mathbb{1})(|\psi_0\rangle\langle\psi_0|)\right) = 0. \tag{B.24}$$

Let us recall that from Proposition 2 we know that

$$||\Phi_U - \Phi_{\mathbb{1}}||_\diamond = 2\sqrt{1-\nu^2\left(U\right)}, \tag{B.25}$$

and from Theorem 2 it holds that

$$||\mathcal{P}_U - \mathcal{P}_{\mathbb{1}}||_\diamond = \min_{E\in\mathcal{DU}(\mathcal{X})} ||\Phi_{UE} - \Phi_1||_\diamond. \tag{B.26}$$

Utilizing Eq. (B.25) and (B.26) we obtain that

$$\max_{E\in\mathcal{DU}(\mathcal{X})} \nu^2\left(UE\right) = 0. \tag{B.27}$$

By the property that $0 \in W_{\sqrt{1-\delta}}(UE)$ whenever $0 \in W(UE)$, we have that

$$\max_{E\in\mathcal{DU}(\mathcal{X})} \nu^2_{\sqrt{1-\delta}}\left(UE\right) = 0. \tag{B.28}$$

Secondly, we consider the situation when $\|\mathcal{P}_U - \mathcal{P}_{\mathbb{1}}\|_\diamond < 2$. Let

$$E_0 \in \arg\max_{E\in\mathcal{DU}(\mathcal{X})} \nu\left(UE\right). \tag{B.29}$$

Again, by referring to Eq. (B.25) and (B.26) we obtain that $\nu\left(UE_0\right) > 0$. Let λ_1, λ_d be a pair of the most distant eigenvalues of UE_0. Note that the following relation holds

$$\nu\left(UE_0\right) = \frac{|\lambda_1 + \lambda_d|}{2}. \tag{B.30}$$

As the assumptions of the Lemma 5 in Appendix B are saturated for the defined E_0, we consider the input state

$$|\psi_0\rangle = \sum_{i=1}^{d} \sqrt{\rho_0}|i\rangle \otimes |i\rangle, \tag{B.31}$$

where the existence of ρ_0 together with its properties are described in Lemma 5 and Lemma 10. Let us define sets

$$\mathcal{C}_i \coloneqq \left\{\Omega : 0 \leq \Omega \leq \mathbb{1},\, \mathrm{tr}\left(\left(\mathbb{1}-\Omega\right)\frac{\sqrt{\rho_0}|i\rangle\langle i|\sqrt{\rho_0}}{\langle i|\rho_0|i\rangle}\right) \leq \delta\right\} \tag{B.32}$$

for each i such that $\langle i|\rho_0|i\rangle \neq 0$. Now we take the measurement Ω_0 as

$$\Omega_0 = \sum_{i=1}^{d} |i\rangle\langle i| \otimes \Omega_i^\top, \tag{B.33}$$

where $\Omega_i \in \mathcal{C}_i$ is defined as

$$\Omega_i \in \arg\min_{\widetilde{\Omega}\in\mathcal{C}_i} \mathrm{tr}\left(\widetilde{\Omega}\frac{\sqrt{\rho_0}U|i\rangle\langle i|U^\dagger\sqrt{\rho_0}}{\langle i|\rho_0|i\rangle}\right) \tag{B.34}$$

for each $i \in \{1,\ldots,d\}$ 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

$$\alpha(\psi_0,\Omega_0) = 1-\mathrm{tr}\left(\Omega_0(\mathcal{P}_{\mathbb{1}}\otimes\mathbb{1})(|\psi_0\rangle\langle\psi_0|)\right) = 1-\sum_{i=1}^{d}\mathrm{tr}\left(\Omega_i\sqrt{\rho_0}|i\rangle\langle i|\sqrt{\rho_0}\right) \leq \delta. \tag{B.35}$$

Hence, it remains to show that for this setting

$$\mathrm{tr}\left(\Omega_0(\mathcal{P}_U\otimes\mathbb{1})(|\psi_0\rangle\langle\psi_0|)\right) = \max_{E\in\mathcal{DU}(\mathcal{X})}\nu^2_{\sqrt{1-\delta}}(UE)\,. \tag{B.36}$$

Direct calculations reveal that

$$\begin{aligned}\mathrm{tr}\left(\Omega_0(\mathcal{P}_U\otimes\mathbb{1})(|\psi_0\rangle\langle\psi_0|)\right) &= \sum_{i=1}^{d}\mathrm{tr}\left(\Omega_i\sqrt{\rho_0}U|i\rangle\langle i|U^\dagger\sqrt{\rho_0}\right)\\ &= \sum_{i=1}^{d}\langle i|\rho_0|i\rangle\mathrm{tr}\left(\Omega_i\frac{\sqrt{\rho_0}U|i\rangle\langle i|U^\dagger\sqrt{\rho_0}}{\langle i|\rho_0|i\rangle}\right).\end{aligned} \tag{B.37}$$

Let us define

$$\beta^{|i} = \mathrm{tr}\left(\Omega_i\frac{\sqrt{\rho_0}U|i\rangle\langle i|U^\dagger\sqrt{\rho_0}}{\langle i|\rho_0|i\rangle}\right). \tag{B.38}$$

Note that due to Lemma 10, the absolute value of the inner product between pure states $\frac{\sqrt{\rho_0}|i\rangle}{\|\sqrt{\rho_0}|i\rangle\|}$ and $\frac{\sqrt{\rho_0}U|i\rangle}{\|\sqrt{\rho_0}|i\rangle\|}$ is the same for every $i \in \{1,\ldots,d\} : \langle i|\rho|i\rangle \neq 0$. Therefore, we can consider the certification of pure states conditioned on the obtained label i with statistical significance δ. From the Proposition 13 we know that $\beta^{|i}$ depends only on such an inner product between the certified states, hence $\beta^{|i} = \beta^{|j}$ for each $i,j : \langle i|\rho|i\rangle, \langle j|\rho|j\rangle \neq 0$. Therefore, the value of $\beta^{|i}$ will depend on $\left|\frac{\lambda_1+\lambda_d}{2}\right|$. Thus, without loss of generality. we can assume that $p^{|1}_{\mathrm{II}} \neq 0$, and hence

$$\sum_{i=1}^{d}\langle i|\rho_0|i\rangle\beta^{|i} = \beta^{|1} = \mathrm{tr}\left(\Omega_1\frac{\sqrt{\rho_0}U|1\rangle\langle 1|U^\dagger\sqrt{\rho_0}}{\langle 1|\rho_0|1\rangle}\right). \tag{B.39}$$

In the remaining part of the proof we will show that

$$\beta^{|1} = \max_{E \in \mathcal{DU}(\mathcal{X})} \nu^2_{\sqrt{1-\delta}}(UE)\,. \tag{B.40}$$

It is sufficient to study two cases depending on the relation between $\sqrt{\delta}$ and the inner product

$$\left|\frac{\langle 1|\rho_0 U|1\rangle}{\langle 1|\rho_0|1\rangle}\right| = \left|\frac{\lambda_1+\lambda_d}{2}\right|. \tag{B.41}$$

In the case when $\left|\frac{\lambda_1+\lambda_d}{2}\right| \leq \sqrt{\delta}$, then, due to Proposition 13, we get $\beta^{|1} = 0$. On the other hand, we know that $0 \in W_{\sqrt{1-\delta}}(UE_0)$ and hence also

$$\max_{E \in \mathcal{DU}(\mathcal{X})} \nu^2_{\sqrt{1-\delta}}(UE) = 0. \tag{B.42}$$

In the case when $\left|\frac{\lambda_1+\lambda_d}{2}\right| > \sqrt{\delta}$, then, from Proposition 13, we know that

$$\beta^{|1} = \left(\left|\frac{\lambda_1+\lambda_d}{2}\right|\sqrt{1-\delta} - \sqrt{1-\left|\frac{\lambda_1+\lambda_d}{2}\right|^2}\sqrt{\delta}\right)^2. \tag{B.43}$$

On the other hand, for $E_0 \in \mathcal{DU}(\mathcal{X})$ satisfying Eq. (B.29), we have

$$\nu^2_{\sqrt{1-\delta}}(UE_0) = \left(\left|\frac{\lambda_1+\lambda_d}{2}\right|\sqrt{1-\delta} - \sqrt{1-\left|\frac{\lambda_1+\lambda_d}{2}\right|^2}\sqrt{\delta}\right)^2. \tag{B.44}$$

By the particular choice of $E_0 \in \mathcal{DU}(\mathcal{X})$, this value is equal to $\max_{E\in\mathcal{DU}(\mathcal{X})} \nu^2_{\sqrt{1-\delta}}(UE)$, hence combining the above equations we finally obtain

$$\beta^{|1} = \max_{E \in \mathcal{DU}(\mathcal{X})} \nu^2_{\sqrt{1-\delta}}(UE)\,. \tag{B.45}$$

To sum up, we indicated Ω_0 and $|\psi_0\rangle$ for which the optimized probability of type II error was equal to $\max_{E\in\mathcal{DU}(\mathcal{X})} \nu^2_{\sqrt{1-\delta}}(UE)$. Combining this with the previously proved inequality

$$\beta_\delta \geq \max_{E \in \mathcal{DU}(\mathcal{X})} \nu^2_{\sqrt{1-\delta}}(UE) \tag{B.46}$$

gives us Eq. (B.16) and proves that the proposed strategy $|\psi_0\rangle, \Omega_0$ is optimal. ■