COMENIUS UNIVERSITY IN BRATISLAVA
FACULTY OF MATHEMATICS, PHYSICS AND INFORMATICS

Labeling outcomes of Quantum Devices and Higher-order distinguishability

Dissertation

July 2024

Mgr. Nidhin Sudarsanan Ragini

COMENIUS UNIVERSITY IN BRATISLAVA

FACULTY OF MATHEMATICS, PHYSICS AND INFORMATICS

Labeling outcomes of Quantum Devices and Higher-order distinguishability

Dissertation

Study program: **Theoretical and Mathematical Physics**
Study field: **13.-Physics**
Training center: **Institute of Physics, Slovak Academy of Sciences**
Supervisor: **Doc. Mgr. Mário Ziman, PhD**

July 2024 **Mgr. Nidhin Sudarsanan Ragini**

36222695

Comenius University Bratislava
Faculty of Mathematics, Physics and Informatics

THESIS ASSIGNMENT

Name and Surname:	Nidhin Sudarsanan Ragini
Study programme:	Theoretical Physics and Mathematical Physics (Single degree study, Ph.D. III. deg., full time form)
Field of Study:	Physics
Type of Thesis:	Dissertation thesis
Language of Thesis:	English
Secondary language:	Slovak

Title: Labeling outcomes of quantum devices and higher-order distinguishability

Annotation: Student will address the distinguishability questions for higher-order quantum structures. In particular, he will focus on the task of outcome labeling of quantum measurements and potentially also the role of incompatibility for the distinguishability of quantum measurements of higher-order structures.

Literature: 1. G. Chiribella, G. M. D'Ariano, and P. Perinotti, Theoretical framework for quantum networks, Phys. Rev. A 80 , 022339 (2009)
2. A. Bisio et al., Quantum Networks: General Theory and Applications, Acta Physica Slovaca 61, 273–390 (2011)
3. Teiko Heinosaari and Mário Ziman, The Mathematical Language of Quantum theory: From Uncertainty to Entanglement, Cambridge Univ. Press, 2012)

Tutor:	doc. Mgr. Mário Ziman, PhD.
Department:	FMFI.KTF - Department of Theoretical Physics
Head of department:	doc. RNDr. Tomáš Blažek, PhD.
Assigned:	09.08.2019
Approved:	09.08.2019

prof. RNDr. Peter Moczo, DrSc.
Guarantor of Study Programme

.... Student

.... Tutor

36222695

Univerzita Komenského v Bratislave
Fakulta matematiky, fyziky a informatiky

ZADANIE ZÁVEREČNEJ PRÁCE

Meno a priezvisko študenta: Nidhin Sudarsanan Ragini
Študijný program: teoretická fyzika a matematická fyzika (Jednoodborové štúdium, doktorandské III. st., denná forma)
Študijný odbor: fyzika
Typ záverečnej práce: dizertačná
Jazyk záverečnej práce: anglický
Sekundárny jazyk: slovenský

Názov: Labeling outcomes of quantum devices and higher-order distinguishability
Označovanie výsledkov kvantových zariadení a rozlišovanie štruktúr vyššieho rádu

Anotácia: Študent sa zameria na otázky rozlíšiteľnosti kvantových štruktúr vyššieho rádu. Konkrétne, na úlohu označenia výsledkov kvantových meraní a potenciálne tiež na úlohu kvantovej nekompatibility pre rozlíšiteľnosť kvantových meraní vyššieho rádu.

Literatúra:
1. G. Chiribella, G. M. D'Ariano, and P. Perinotti, Theoretical framework for quantum networks, Phys. Rev. A 80 , 022339 (2009)
2. A. Bisio et al., Quantum Networks: General Theory and Applications, Acta Physica Slovaca 61, 273–390 (2011)
3. Teiko Heinosaari and Mário Ziman, The Mathematical Language of Quantum theory: From Uncertainty to Entanglement, Cambridge Univ. Press, 2012)

Školiteľ: doc. Mgr. Mário Ziman, PhD.
Katedra: FMFI.KTF - Katedra teoretickej fyziky
Vedúci katedry: doc. RNDr. Tomáš Blažek, PhD.

Dátum zadania: 09.08.2019

Dátum schválenia: 09.08.2019

prof. RNDr. Peter Moczo, DrSc.
garant študijného programu

....

študent školiteľ

Dedicated to Achan, Amma, Aniyathi, Praseeda and my dearest friends.

Acknowledgements

Foremost, I thank my supervisor, Mario Ziman, and Research Center for Quantum Information (RCQI), Institute of Physics, Slovak academy of sciences for giving me an opportunity, time and space for probing fundamental limitations of quantum theory imposed on itself through its mathematical structures. I also thank Mario for his support during my period there. I started my doctoral studies more or less in tandem with a global pandemic; As most of the people around the globe would agree that those were periods of confusion and distress, I also shared a version of it: To be in a foreign nation under those conditions and with near to no real life interactions with people, my emotional well being was in a turmoil. I thank Ashin, who was probably the only person I was able to converse in real life and feel human during that time. He was also a good friend and a mentor for me, in a couple things, including cooking. I am thankful to myself that I came out of the ordeal. I also cherish gratefully about the post-pandemic trip I and my closest friend Arjun had, which revived myself to a more alive version.

I am grateful to all of my colleagues from RCQI and QUTE those who created a friendly environment in general and from those I learnt a fair share of things, academic and otherwise. I remember Maria, Pavol, Roman, Andrej, Peter, Daniel, David, Hamed, Gilberto, Younes and Leevi. I especially thank Arash, Natalia and Fereshte for their kindness and support. I thank all the researchers and academics I had the previlige of meeting and interacting with, those who have enriched my academic as well critical thinking lexicon.

I thank my peers, Nana, Tony, Denda and Jaro for creating a conducive office for work as well as for making the air more lively. Nana, also being a good friend, shared a fair share of same technical hurdles as mine, throughout the past years here and was someone I could always whine with. Tony was godsend that at last I was able to speak my mother tongue and feel little bit close to home; He was a good friend and support.

I cannot see my doctoral studies coming to a fruition without the academic support and nourishment by Arash and Sazim (Da); Both of them have supported me in various capacities, through, but not restricted to, stimulating discussions, collaborations and mentoring. Moreover, Da and his partner Amina (Didi), were a constant source of moral support, financial support and love; Especially during times of crisis, Didi has had always come up with the most grounded suggestions and plans of action. I already miss the evening dinners we used to have at Da and Didi's place and the discussions we used to have regarding anything and everything. I write this as being a fan of Didi's biriyani!

I thank Arash, Ardra, Samgeeth and Sazim for meticuluously going through substantial parts of the thesis and commenting on various aspects as well as pointing out errata. With their help, this thesis is more readable, more insightful and with lesser errors.

I cannot express my gratitude enough for my closest friends who are elsewhere in Europe, Arjun, Aarya, Abhi, Sukritha, Karthika. The gratitude extends to my closest in Bratislava, including Diya and Balu. They were a constant support and were there for me during different times of need, irrespective whether it was emotional or financial; Without doubt, I can say that acknowledging properly would require couple of pages for each. Aarya, who was also writing her doctoral thesis in Italy, in tandem with me, related greatly with me and always found time to check on me. I seriously doubt whether I was able to be a better friend to mine during writing of this thesis. Same gratitude goes for the friends I made lately in RCQI, Samgeeth, Ardra, Soham, Ieline and Sal, for their unwavering as well as diverse supports. Apart from RCQI, the list continues to include Ijas, Sylvia and Pranav, Bhavana and Gokul; Especially Sylvia, who was a constant support and confidence booster for me. For sure, I am not missing out on you guys, Vishakhettan, Ashiq and Kiranettan; we will hang out more often. To wind up, my life here in Bratislava, and in general Europe, would have been of lesser meaning and faded colours without you people in it, and I shall always cherish the memories I share with each of you and love you. I should also not forget my friends back home whom, even though we talk with the lowest frequencies, have always been there to remember and support me; Gopika, Rashid, Abhishek, Gokul and Sree.

My later life in Bratislava cannot go by without mentioning the restaurant "Dosa and Samosa (D&S)", which is run by my closest friend Balu. Talking about D&S, I cannot go by without mentioning another close friend of mine, Gokul and our evening hang outs, with Balu, and discussions at D&S. Our hang outs helped me decompress myself after a day's stressful mental gymnastics at work. Along with this, the delicious dishes prepared by Pandey Bhaiyya satiated a long day's hunger as well as my heart. The friends I made there and the different people with whom I struck conversations enriched my "social intelligence" to some capacity.

The last paragraph is usually reserved for family and understand well enough why. Without support of my family, including my Achan (father), my Amma (mother), my Aniyathi (sister) and my partner Praseeda, this whole journey which spanned almost half a decade would have been insufferable without the thought that you people are constantly backing me up, rooting for me and loving me, even though all of you guys were miles apart from me. Culturally, thanking my parents would present itself as awkward but still I take this opportunity thank them for their, as well as my sister's, support and prayers, and the forces those have assisted me. I thank Praseeda for prioritising and finding time to talk to me during her intense civil service exam preparation. She

has always supported me to come down to reality whenever my mind takes on bungee jumping. I don't know whether I was able to be a better son, brother and partner during the time of thesis writing, but I am grateful for their immense understanding. Love you three-thousand!

Abstrakt

V tejto práci definujeme úlohu spracovania kvantovej informácie, ktorú označujeme ako "kvantové označovanie", a identifikujeme ju ako špecifický prípad úloh rozlišovania týkajúcich sa kvantových pozorovateľných veličín alebo kladne-semidefinitných operátorových mier. Tento špecifický prípad vzniká vtedy, keď súbor, ktorý požadujeme rozlíšiť, pozostáva zo všetkých možných pozorovateľných veličín, ktoré majú identické, ale permutované efekty k výsledkom. Tvrdí sa, že tieto úlohy majú význam aj vtedy, keď sa asociácie medzi efektmi a výsledkami kvantového meracieho zariadenia nejakým spôsobom stratia, čím sa stane ,,neoznačeným". Túto úlohu skúmame v rámci matematických štruktúr vyššieho rádu, ktorá zahŕňa najvšeobecnejšie možné scenáre. Následne skúmame úlohu v rámci režimu jedného a viacerých použití a pre rôzne možné schémy vrátane perfektnej, jednoznačnej a minimálnej chyby.

Štatistický pojem kvantovej nekompatibility a jej "zdrojová" úloha pri kvantovom spracovaní informácie boli predmetom intenzívneho štúdia v posledných rokoch. Tento pojem súvisiaci s pojmom spoločnej merateľnosti, resp. realizovateľnosti kvantových zariadení, bol rozšírený na kvantovo-teoretické merania vyšších rádov, označované ako kvantové testery. Ako druhú tému v tejto práci skúmame nekompatibilitu kvantovo-teoretických meraní vyššieho rádu v úlohách rozlišovania.

Kľúčové slová: kvantová diskriminácia, kvantové hrebene, kvantové testery, kvantová teória vyššieho rádu, POVM, kvantové merania

Abstract

In this thesis, we introduce a quantum information processing task which we refer to as "quantum labeling" and identify that it as a specific instance of distinguishability tasks concerning quantum observables or positive-operator-valued measures. This specific instance arises when the ensemble we require to distinguish is composed of all possible observables having identical but permuted associations of composing effects to the outcome labels. These tasks are also claimed to be of relevance when the associations between effects and outcome labels of a quantum measurement device is lost somehow, rendering it "unlabeled". We investigate this task using higher-order quantum theoretic notions so as to consider the full extend of possibilities. Subsequently, we study the task within the single and multiple-shot regimes and for different possible schemes including perfect, imperfect unambiguous and minimum-error.

The statistical notion of quantum incompatibility and its role as a resource in various quantum information processing tasks have been studied in the recent years. This notion related to joint measureability or implementability of quantum devices has been extended to higher-order quantum theoretic measurements, referred to as quantum testers. In this thesis, as a second avenue other than that on labeling tasks, we explore the resourcefulness of incompatibility of higher-order quantum theoretic measurements in distinguishability tasks.

Keywords: quantum discrimination, quantum combs, quantum testers, higher-order quantum theory, POVM, quantum observable.

Contents

Chapter 1

Introduction

The encompassing theme of this thesis is centred around higher-order quantum theoretic notions and the associated framework, which was developed through the pioneering works done in the late 2000's. In this introduction, we start by briefly discussing the central notions of this framework, intertwined with the associated historical developments. This thesis encompasses two distinct avenues of theoretical research:

- "Quantum labeling tasks"; In this avenue, we primarily use higher-order theoretic notions as mathematical tools for the modeling of accompanied problems and subsequent investigations. [1]
- "Higher-order incompatibility as a resource"; In this avenue, we identify the exclusive resourcefulness possessed by incompatible higher-order quantum theoretic measurements in distinguishability tasks concerning quantum networks.

In order to set the stage for discussing first of the above avenues, we brief about the fundamental relevance of quantum distinguishability tasks in broader areas of quantum science and technology. We essentially discuss how higher-order theoretic tools play their roles in distinguishability tasks and assist our investigation into quantum "labelability". After verbally defining labeling tasks, we steer towards discussing the second avenue of our investigation regarding higher-order quantum incompatibility and its resourcefulness in distinguishability. We conclude this introduction with outlining the rest of the thesis.

Higher-order quantum theory

In the current era of intense scientific research fueled by the desire to achieve and perfect next generation quantum technologies, one of the major theoretical

[1] Despite the usage of these tools it is not exclusive by any means.

tasks the working scientific community wants to achieve is to optimise multi-component quantum networks or circuits, those which are designed to perform specific quantum information processing tasks or quantum computation tasks. These optimisation investigations grow in complexity whenever we start to deal with each and every component of the accompanying networks and as the size of these network grows. As such, the optimisations are streamlined when we envision and mathematically describe the quantum networks as single optimisable mathematical entities. These single descriptors should also be accompanied by prescriptions for physical realisations of these networks, so that they do not render themeselves to be theoretical artefacts. Higher-order quantum theory (HOQT) [CDP09] was developed for this exact purpose and has also been growing so as to reveal fundamental features of quantum theory as well. In fact, in this thesis we can see these two facets of this framework; In our first avenue, we use the framework for the optimisation capabilities it endows and in the second avenue, we learn about a qualitative feature of quantum theory by working within the framework.

HOQT identifies the aforementioned quantum networks with multiple time-step quantum causal quantum processes and associated measurements of these processes and duly assigns single object mathematical descriptors to them[2]. The development of this framework was initiated by Chiribella, D'Ariano and Perinotti in their seminal work [CDP08c]. In this work they introduced the notion of "quantum combs"[3], which were identified as single descriptors of transformations of multiple quantum operations spread across in time. As such, combs describe multiple time-step causal quantum circuits or networks with open slots, where quantum operations can be "plugged" into the slots and thus resulting in a transformed quantum transformation. In fact, rather than just quantum operations, legitimate quantum combs can themselves be plugged into more general combs to have resultant combs. Moreover, similar to how, in standard quantum theory, a quantum state describes equivalent preparations of a system, a particular quantum comb describes equivalent open quantum networks those result in identical output transformations, given identical input transformations. Digressing from the discussion, we would expect that all such quantum process transformations to be "causal", as in to have a definite causal ordering of the component transformations, but we are bound to use that qualifier since Oreshkov, Costa and Časlav kickstarted a parallel research avenue with their work [OCB12], leading to the

[2]This means single time-step processes and assiciated notions are a special case; As such, states, single time-step quantum channels and quantum observables are also described by this framework. In fact they are the building blocks of it.

[3]The choice of the word "comb" is due to the fact that these circuits with open slots resemble a hair comb.

"process matrix framework" wherein processes with indefinite causal ordering are accounted for and studied as legitimate objects and the higher-order theory of quantum combs can be considered as a specific instance in that framework where there is definite causal ordering. For combs, their causal structure translate itself as a set of recursive normalisation conditions. Returning to our discussion, in [CDP08c] it was also identified that a quantum comb with N open slots[4], having $N+1$ time steps, is equivalent to a composition of $N+1$ memory channels [KW05].

Measurements or testing procedures of quantum combs, which result in probability distributions, are described by the notion of quantum testers; Similar to combs, testers are also quantum networks with open slots and "combining" a comb and an appropriate tester, for the testing procedure, results in a closed circuit which outputs classical information. They were introduced in the subsequent work [CDP08d] by Chiribella, D'Ariano and Perinotti, where they investigated the general binary distinguishability tasks for quantum combs and quantum testers described the experimental networks distinguishing the combs. Since testers are also networks with open slots, they are also charaterised by the number of slots they possess. To test a quantum comb with N open slots, it should be combined with testers with $N+1$ open slots. Moreover, similar to quantum observables, testers also satisfy normalisation conditions but peculiar and non-unique. On top of this normalisation, there exists the set of recursive normalisation conditions which reflect the causal structure of the tester network, similar to those possessed by combs. Simultaneously with the above work by Chiribella *et.al.*, Ziman also introduced and characterised the notion of "process POVMs" which described general testing networks designed to test single time-step quantum operations in [Zim08]. In the language of Chiribella, D'Ariano and Perinotti, process POVMs translate to quantum testers with a single open slot. As a side comment, the notions quantum combs and testers attempt to describe were equivalently described by Gutoski and Watrous in their work [GW07] where they studied the networks as quantum games.

To showcase the growing relavance of HOQT, we can look the following non-exhaustive list of areas and quantum information processing tasks where the framework has been used:

- Optimisation of networks designed to clone unitary transformations [CDP08b].
- Optimisation of networks designed to store-and-retrieve or "quantum learn" unitary transformations [BCD+09a, SBZ19, PZ22].

[4] N being a natural number.

– Design and optimisation of networks implementing programmed functions of the given unknown unitary transformation [BCD+09b]. A related network of interest is that inverts unknown transformations [YSM23].

– In quantum metrology, for design and optimisation of networks for parameter estimation [Yan19, LHYY23, BLBSM23, KDM+24].

– Design of networks those can realise desired evolutions by only having access to a restricted set of evolutions [OKSM24].

– It has been used to study distinguishability of quantum memory channels [CDP08d, CDP09] and wherein they showed that memory effects ramified from combs are crucial for achieving optimal distinguishability. The framework has been extensively used to study distinguishability, and other quantum hypothesis testing tasks, of channels [SZ09, ZS10, Jen21, BMQ22, BMQ21, KM24]and measurements [SZ14].

– The mathematical structures of HOQT notions themselves have been studied and characterised [DPS11, Jen12b, Jen12a, Jen13].

Quantum distinguishability tasks

Quantum distinguishability tasks are based on, or reveal, the quantum feature of distinguishability or indistingusihability of a given finite collection of quantum devices, those which can be considered as belonging to the same type. In an instance of such a task, provided some non-trivial prior information regarding the instance, our goal is to identify or equivalently distinguish an unknown device, occuring from the collection, from the rest of the devices. As such, one of the major goals of theoretical investigation of these tasks is to find optimal experiments or test procedures which result in the maximal distinguishability possible for a given collection of devices; These quantum optimisation problems, in general, result in optimisations over convex sets of test procedures. Moreover, these tasks are instances of general quantum hypothesis testing tasks, those wherein we test finite collections of hypotheses made about some quantum aspect by designing and performing experiments.

Study of these tasks was initiated by C.W. Helmstrom in his seminal work of 1969 [Hel69], wherein which he formulated and studied the binary distinguishability task for quantum states. It was discovered that perfect distinguishability of a collection of states is equivalent to their mutual orthogonality; Consequently, the maximal number of states that can be perfectly distinguished is equal to the dimension of the system under study. This orthogonality/non-orthogonality or, more generally, distinguishability poses distinguishability tasks as one of the fundamental classes of tasks that has pragmatic, as well as foundational significance in quantum science, spanning from quantum computation, quantum cryptography, quantum communication, quantum information theory, quantum entanglement, to quantum cloning. Moreover, studies on the binary versions these tasks lead to quantitative similarity measures, for the type of quantum objects under study, with the operational significance related to the task. One such measure appearing in case of quantum states is captured by the celebrated Helstrom-Holevo bound [Hel69, Hol98]. Even after more than 50 years of inception, investigations on state distinguishability problems are still active in the quantum foundations and information theory research communities. Some of these investigations reveal the crucial role distinguishability plays in relation to seemingly distant notions within quantum information science. To cite two such explorations, we can look at [BNV13], where Brunner *et.al.* constructed a device-independent dimension witness [BPA+08]; This construction was based on the fact that the maximal number of perfectly distinguishable states, with single copy accessibility, is equal to the dimension of the system. Another exploration is where state distinguishability is used in constructing quantum observable incompatibility witnesses [CHT19].

Distinguishability tasks were then studied for transformations of states, modeled by quantum operations. In binary tasks for quantum operations, we would send known quantum states through the selected unknown operation and then measure this state so as to make conclusion whether the implemented operation is this or that. Consequently, distinguishing such devices essentially boils down to distinguishing the transformed states. During the initial studies, one crucial difference from general state distinguishability tasks that adds to the richness of operation distinguishability tasks, discovered by Sacchi [Sac05], is the fact that for some collections of operations, quantum entanglement was found to be a resource in improving the distinguishability. As such, these two aspects of choosing the best test state as well as the assistance by auxiliary quantum systems led to the operational distance given by the diamond distance or completely bounded distance of two operations.

Eventually, distinguishability tasks of quantum measurements, when they are described by normalised-positive-operator-valued measures [Hol11, HZ11], were studied [JFDY06, ZH08, Sed10, SZ14, PPKK18, KPP20, DBSA21, PPK+21]. Essentially, all observables can be considered as quantum operations when they are modeled as quantum-to-classical, measure-and-prepare channels [Hol98]. Consequently every observable discrimination problem can be recast into a quantum operation discrimination problem. A shift in this approach began with the work [JFDY06], where Ji *et.al.* exploited the structure of the involved observables to effectively identify each of them, through measuring the unknown observable on carefully chosen states. Such schemes are referred to as simple schemes in this thesis.

Now, distinguishability tasks can be viewed as either belonging to the so-called single-shot or multiple-shot regime, based on whether we have access to a single copy/use/query of the devices or more, respectively. It was found that, for state ditinguishability tasks, having access to more number of copies of state could result in better distinguishability [ACMnT$^+$07, CMnTM$^+$08]. This has implications to all other distinguishability tasks since measurements of states are to these tasks, irrespective of whether explicitly identifiable or not. On top of this, it was also found that multiple-shot schemes are better in quantum operation distinguishability [CDP08d, HHLW10, Chi12], when other quantum effects including memory ones can result in improving the task.

As we discussed earlier, since quantum states, operations and measurements can be identified as special instances of quantum combs, we can in general consider the distinguishability tasks of these devices as special instances of quantum comb distinguishability tasks. Aligning with this identification, in chapter 3 we discuss the aspects of distinguishability tasks through these most general cases and then considering the relevant expositions as special cases of the same.

Quantum labeling tasks

A quantum measurement[5], when described as a normalised-positive-operator-valued measure (POVM) or observable[6], is characterised by a set of outcome labels, a set of effects which is a resolution of the identity operator and a set of prescriptions which assign or associate these outcome labels to the corresponding effects. It could happen that we have access to a measurement device implementing such an observable and are oblivious regarding the aforementioned prescriptions. We can consider this observable as "unlabeled" since we are unaware to what measurement events the outcome labels correspond to. In [ZH08], Ziman and Heinosaari refered to this notion as to correspond to equivalence classes of observables, whereby such a class includes all those observables with identical range, while taking into account possible multiplicities of effects. Subsequently, in [ZHS09], Ziman *et.al* introduced the notion of an "unlabeled observable" as giving rise to an equivalent class of "labeled" observables in the following sense: Two labeled observables are equivalent to each other, if they have permuted as-

[5]In this thesis, we restrict our studies to quantum measurements on finite-dimensional quantum systems with finite and discrete outcome sets.

[6]In this thesis, as it will be elaborated in the next chapter, POVMs and observables are used to describe the same notion and are synonymous with each other.

sociations of outcome labels to the same collection of effects. As such, when we are given an observable from such a class, the information we possess is that it could be one among the possible labeled observables. In such scenarios, we can task ourselves to identify the label-effect associations, given in prior that we have knowledge regarding the composing effects, essentially identifying the labeled observable. In reality, we would want to perform this task due to a variety of circumstances, ranging from sabotaged labels to having distrust on the manufacturer of the measurement device that we wish to certify it by ourselves. Motivated to perform this task, with finite uses or queries of the unlabeled measurement device, in this thesis, we introduce the task of "quantum labeling".

Prior to investigating these tasks, we identify that these tasks are equivalent to a specific class of distinguishability tasks of observables, where within each collection of observables to be distinguished, each and every observable is described by the same set of effects but possesses permuted associations. In other words, given an unlabeled observable, the labeling task is a distinguishability task concerning the corresponding equivalent class of labeled observables where each of them occur with equal probabilities.

Our investigation on labeling tasks is partitioned into two chapters. The first chapter, chapter 4, the studies are restricted to the single-shot regime. In this chapter, we identify classes of two-outcome observables those can be perfectly labeled, introduce an operational distance for "labelability" and arrive at few no-go theorems for two-outcome as well multiple-outcome observables. In the second chapter on labeling tasks, chapter 5, we investigate the task within the multiple-shot regime, so as to probe whether adaptive networks and associated quantum memory effects result in improving labelability.

Higher-order incompatibility and its resourcefulness

Similar to quantum distinguishability being a prime and genuine feature of quantum theory, incompatibility or joint-measurability/joint-implementability is another such feature. It conveys the fact that there exist collections of quantum devices those can be implemented jointly by a single device. More crucially, there exist collections which do not admit such joint devices. Historically, this notion was exclusively restricted to quantum measurements of physical quantities described by self-adjoint operators and existence of joint measurement of such a collection of operators was synonymous with their mutual commutativity. When quantum measurements were modeled as positive-operator-valued-measures (POVMs), this ramification of joint-measureability as commutativity was rendered to being a special and specific case [LP97, Lah03].

Similar to how quantum state entanglement has been identified as a state resource in a plethora of quantum information processing tasks and games (being necessary for two-party Bell violations, quantum steering, quantum teleportation, quantum key distribution), incompatibility of quantum measurements has been identified as a measurement resource. It is also a necessary resource for Bell violations [Fin82, MAG06, WPGF09], quantum steering [QVB14, UMG14] and for one-sided quantum key distribution [BCW+12]. A crucial finding, from the point of view of this thesis, was discovered by Skrzypczyk, Šupić and Cavalcanti that all collections of incompatible POVMs are resourceful for a particular two-party state distinguishability game they investigated [SŠC19]. In other words, each and every collection of incompatible POVMs outperform any compatible collection in this game. The aforementioned work established that POVM incompatibility is a resource for state distinguishability.

We started this discussion by referring to the notion of joint measurability or implementability for collections of general quantum devices but afterwards we exclusively discussed incompatibility dealing with measurements of states. The notion was extended to describe joint-implementability of quantum channels [HM17], quantum instruments [MF22, LS24, JC24] and pairs of channels and POVMs [HRRZ18]. But more relevant for this thesis, the notion was extended to quantum testers by Sedlák *et.al.* in [SRCZ16]. This incompatibility of these higher-order quantum theoretic measurements has distinct features when compared to POVM incompatibility; One contrasting fact being that collections of testers composed of mutually commuting operators necessarily do not need to be jointly implementable [SRCZ16]. Inspired by the above works, we investigate in this thesis so as to reveal whether this higher-order incompatibility of testers is a resource in comb distinguishability tasks. For this purpose, we introduce a comb distinguishability game, a generalised version of the game found in [SŠC19], and investigate it. This investigation and the associated result which finds that "tester incompatibility is indeed a resource in comb distinguishability" forms chapter 6. In the wake of this finding, we can identify that ours is a generalised result having the finding in [SŠC19] as a special case.

1.1 Publications and preprints

Publications and preprints those which came out of the doctoral studies and part of this thesis are enlisted here.

- Nidhin Sudarsanan Ragini, Mário Ziman
  Single-shot labeling of quantum observables
  Phys. Rev. A 109, 052415, (2024), arXiv:2407.05351

- Seyed Arash Ghoreishi, Nidhin Sudarsanan Ragini, Mário Ziman, Sk Sazim
  Multiple-shot labeling of quantum observables
  arXiv:2407.05392

- Nidhin Sudarsanan Ragini, Sk Sazim
  Higher-order incompatibility improves distinguishability of causal quantum networks
  arXiv:2405.20080

Chapter 2

From Preliminaries to Higher-order quantum theory

In this chapter, we discuss skeletal mathematical notions based on which the "standard" quantum theoretic framework, for systems with finite quantum degrees of freedom, is built upon. Moreover, we discuss relevant mathematical tools required for this thesis, one of such discussions is a brief one on semi-definite programming, which is a central tool used in chapter 6. After discussing these preliminaries, we sketch and discuss notions forming the framework of quantum theory within a scope that is relevant for the thesis. Subsequently, we discuss and sketch basic higher-order quantum theoretic notions; Again to a degree and depth proportional to their relevance for this work.

2.1 Mathematical preliminaries

As already mentioned, in this work, the studies are carried out within the framework of what we call standard quantum theory, which is constructed on Hilbert spaces and associated structures. This framework's mathematically rigourous development was initiated by John von Neumann in his series of works from late 1920's, those were then amalgamated into his book, *Mathematical foundations of quantum mechanics* (See [vNB18] for the new edition). This framework is based on Hilbert spaces, linear operators between such spaces and related and induced structures on these spaces. Before discussing this framework and ramified notions of thereof, those are relevant for the thesis, in the next section, in this section we discuss adequeate mathamatical notions which are relevant for the discussion as well as the thesis. As suggested earlier, since linear operators on Hilbert spaces are central to the framework, we begin by discussing them and relevant, associated notions.

2.1.1 Hilbert space structures and operators

We establish that each and every linear vector space underlying any of the notions used and developed in this thesis is restricted to be finite-dimensional, unless stated otherwise; They are also assumed be on complex fields $\mathbb{C}$. That is, if $d = \mathbf{dim}(\mathcal{V}_{\mathbb{C}})$ is the dimension of the vector space $\mathcal{V}_{\mathbb{C}}$, then $d < \infty$ and $\mathcal{V}_{\mathbb{C}}$ is isomorphic to the space $\mathbb{C}^d$ [HZ11]. Vectors of this space are column vectors $\varphi = (x_1, \ldots, x_d)^\top$ with x_j being complex numbers and "$\top$" corresponding to the matrix transposition operation and is performed with respect to a chosen basis. Deploying the Dirac bra-and-ket notation, these vectors are represented as "kets" $|\varphi\rangle$; The dual vector of such a vector is represented using the corresponding "bra" $\langle\varphi|$, which is arrived at by transposing and conjugating the associated ket vector. With this consensus, we can continue the following discussions.

Inner-products and Hilbert spaces

An inner-product f is a sesqi-linear binary functional on a vector space $\mathcal{V}_{\mathbb{C}}$, $f : \mathcal{V}_{\mathbb{C}} \times \mathcal{V}_{\mathbb{C}} \to \mathbb{C}$, which is conjugate symmetric and positive definite. Any such finite-dimensional vector space equipped with an inner-product is referred to as a (finite-dimensional) Hilbert space[1]. Given the Hilbert space $\mathbb{C}^d$, the chosen inner product between two vectors $|\varphi\rangle = (x_1, \ldots, x_d)^\top$, $|\psi\rangle = (y_1, \ldots, y_d)^\top$ is defined through,

$$\langle\varphi|\psi\rangle = f(|\varphi\rangle, |\psi\rangle) := \sum_{k=1}^{d} x_k^* y_k. \tag{2.1}$$

Here, "$*$" corresponds to the conjugation operation on complex numbers. Equipment of an inner product gives rise to the notion of orthogonality between vectors; Two non-zero vectors, $|\varphi\rangle$ and $|\psi\rangle$, are said to be orthogonal when the inner product of these vanishes, that is, $\langle\varphi|\psi\rangle = 0$. We denote this orthogonality as $|\varphi\rangle \perp |\psi\rangle$. This enables us to have the notion of orthogonal subspaces, wherein all the vectors belonging to two subspaces are mutually orthogonal to each other. Moreover, we can also have orthogonal bases, wherein the basis elements are mutually orthogonal to each other. More specifically, we can have orthonormal bases $\{|\varphi_k\rangle\}_{k=1}^d$ those satisfy $\langle\varphi_i|\varphi_j\rangle = \delta_{ij}$ for all i, j. A specific choice of orthonormal basis we use in the thesis is represented as $\{|1\rangle, \ldots, |d\rangle\}$, where $|k\rangle$ is a column vector with its k^{th} row entry being 1 and rest of the entries being 0.

[1] In fact, inner product spaces those are Cauchy complete with respect to the norm induced by the inner product are referred to as Hilbert spaces. We do not explicitly talk about completeness because all finite-dimensional inner product spaces are Cauchy complete by construction, and we deal only with finite-dimensional spaces in this work.

Linear operators

A linear operator X from a Hilbert space $\mathcal{H}$ to another $\mathcal{K}$ is a map $X : \mathcal{H} \to \mathcal{K}$ which preserves the linear structure of the spaces[2]; If the input (domain) and output (codomain) spaces coincide, then we say X is an operator on that space. Any linear operator can be constructed through "outer-products" of a bra and a ket as $|\psi\rangle\langle\varphi|$, with $|\varphi\rangle \in \mathcal{H}, |\psi\rangle \in \mathcal{K}$, and linear combinations of such products. Now, given any such operator $X : \mathcal{H} \to \mathcal{K}$, by choosing a basis $\{|x_i\rangle\}_{i=1}^{m}$ for $\mathcal{H}$ and $\{|y_j\rangle\}_{j=1}^{n}$ for $\mathcal{K}$, X can be associated with the an $n \times m$ matrix with entries $X(j,i) = \langle j|X|i\rangle$. Here, we discuss few notions regarding operators, which shall be relevant for us.

- ***Eigensystem.*** Given an operator $X : \mathcal{H} \to \mathcal{H}$, a complex number x, which satisfies $X|\varphi_x\rangle = x|\varphi_x\rangle$ for some vector $|\varphi_x\rangle$, is referred to as an eigenvalue of X and $|\varphi_x\rangle$ is said to be the associated eigenvector. The set of all possible pairs of eigenvalues and associated eigenvectors, $\{(x, |\varphi_x\rangle)\}$, is referred to as the eigensystem of the operator X.
- ***Kernel.*** Kernel or null space of an operator X is the subspace of vectors which are mapped to the zero vector $\mathbf{0}$ by X. That is, it is given by $\mathbf{ker}(X) = \{|\varphi\rangle \in \mathcal{H} : X|\varphi\rangle = \mathbf{0}\}$.
- ***Support.*** Support of an operator X is the subspace of vectors which are not mapped to the zero vector $\mathbf{0}$ by X. In other words, support is an orthogonal subspace to the kernel and is given by $\mathbf{supp}(X) = \{|\varphi\rangle \in \mathcal{H} : |\varphi\rangle \perp |\psi\rangle \text{ for all } |\psi\rangle \in \mathbf{ker}(X)\}$. Given two operators X and Y, we denote it by $X \perp Y$ when these two operators have orthogonal supports.
- ***Rank.*** Rank of an operator X is the dimension of its support. That is, $\mathbf{rank}(X) = \mathbf{dim}(\mathbf{supp}(X))$. Such an operator is said to be "full-rank" when $\mathbf{rank}(X) = \mathbf{dim}(\mathcal{H})$. We refer X as being "rank-deficient" when $\mathbf{rank}(X) < \mathbf{dim}(\mathcal{H})$. We can note that such operators are non-invertible.

We denote $\mathcal{L}(\mathcal{H})$ as the set of linear operators on $\mathcal{H}$ and $\mathbf{dim}(\mathcal{L}(\mathcal{H})) = \mathbf{dim}(\mathcal{H})^2$. In fact, this set can be endowed with a linear space structure as well as an inner-product. That is, $\mathcal{L}(\mathcal{H})$ is itself a Hilbert space with the Hilbert-Schmidt inner-product. This inner-product for two operators X and Y is given by $\mathrm{Tr}\left\{X^\dagger Y\right\}$. Here, the "daggered" X corresponds to the adjoint operator of X, which is prescribed through the satisfaction of $\langle\varphi|X\psi\rangle = \left\langle\varphi X^\dagger\middle|\psi\right\rangle$ for all vectors $|\varphi\rangle$ and $|\psi\rangle$.

[2] In the literature, "linear transformation" is used to refer when the domain and codomain spaces are in general different and for the special case when they coincide "linear operator" is used, so as to have a distinction. In this thesis, we refer both of these scenarios to "(linear) operators" and the corresponding spaces should be clear from the context.

In the matrix correspondence, the adjoint of a matrix is arrived by entry-wise conjuagting the matrix and transposing it. The map "Tr" is the so-called trace functional whose action can be defined through what it does to an outer-product, $\mathrm{Tr}\{|x\rangle\langle y|\} = \langle y|x\rangle$. In the matrix correspondence, its action is equivalent to adding up all of the diagonal entries of the operator matrix. Since, sum of diagonal entries correspond to the sum of the eigenvalues, the trace evalues exactly this quantity.

Relevant operators

Here, we briefly discuss some classes of operators which are implicitly or explicitly relevant for this thesis. Let $X \in \mathcal{L}(\mathcal{H})$.

- ***Self-adjoint operators.*** X is self-adjoint or Hermitian if it coincides with its adjoint operator $X^\dagger$, that is, $X = X^\dagger$. Every self-adjoint operator X admits a unique decomposition, referred to as its "spectral decomposition" which is given by
$$X = \sum_{k=1}^{d} \omega_x |\varphi_x\rangle\langle\varphi_x| . \tag{2.2}$$
Here, $(\{\omega_x\}_x, \{|\varphi_x\rangle\}_x)$ is the eigensystem of X with $\{|\varphi_x\rangle\}_x$ being an orthonormal basis for the space $\mathcal{H}$. Self-adjoint operators have real eigenvalues.

- ***Positive definite and positive semi-definite operators.*** X is "positive definite" if $\langle\psi|X\psi\rangle > 0$ for all non-zero $\psi \in \mathcal{H}$. It is "positive semi-definite" if $\langle\psi|X\psi\rangle \geq 0$ for all non-zero $\psi \in \mathcal{H}$. Eigenvalues of a postive definite operator are all positive and since there are no zero eigenvalues, each and every such operator is invertible. For the case of positive semidefinite operators, their eigenvalues are all non-negative and for those operators possessing zero eigenvalues they are rendered non-invertible. In the thesis, when we refer to positive operators, we refer to the later positive semi-definite operators. The set of these operators is denoted $\mathcal{L}_+(\mathcal{H})$. Moreover, the notion of positive operators enables us equip the set of self-adjoint operators with a partial order: Given two self-adjoint operators X and Y, we write $X \geq Y$ when $X - Y$ is a positive operator.

- ***Projectors.*** An operator X is a projector or projection operator if $X^2 = X$. We can note that all outer-products of normalised vectors are projectors, $|\varphi\rangle\langle\varphi|^2 = |\varphi\rangle\langle\varphi|$.

- ***Zero and Identity operators.*** Zero operator on $\mathcal{H}$, denoted by $O \in \mathcal{L}(\mathcal{H})$, maps each and every vector in $\mathcal{H}$ to the zero vector $\mathbf{0} \in \mathcal{H}$, that is,

$O\psi = \mathbf{0}$ for all $\psi \in \mathcal{H}$. Identity operator on $\mathcal{H}$, denoted by $\mathbf{id}_{\mathcal{H}} \in \mathcal{L}(\mathcal{H})$, maps each and every vector in $\mathcal{H}$ to itself, that is, $\mathbf{id}_{\mathcal{H}}\psi = \psi$ for all $\psi \in \mathcal{H}$. We can note that the null space of O is the entire space $\mathcal{H}$, whereas the support of $\mathbf{id}_{\mathcal{H}}$ is the entire space, that it is the complete projector.

– ***Unitary operators.*** An operator X is unitary if it's adjoint coincides with its inverse, that is, $XX^\dagger = X^\dagger X = \mathbf{id}$.

We are now equipped to present a proposition that we shall be using for our analyses.

Proposition 2.1.1. Let $X, Y \in \mathcal{L}(\mathcal{H})$, then $\mathrm{Tr}\{XY\} = 0$ is possible *if and only if* $X \perp Y$, that is they have orthogonal supports. Moreover, if $X, Y \in \mathcal{L}_+(\mathcal{H})$, then $\mathrm{Tr}\{XY\} = 0$ implies $XY = O$.

Tensored Hilbert spaces and related notions

Given two Hilbert spaces, $\mathcal{H}_1$ and $\mathcal{H}_2$, we can construct a composite space $\mathcal{H}_1 \otimes \mathcal{H}_2$, referred to as the tensor product space of the corresponding spaces, with $\mathbf{dim}(\mathcal{H}_1 \otimes \mathcal{H}_2) = \mathbf{dim}(\mathcal{H}_1) \times \mathbf{dim}(\mathcal{H}_2)$. This composite space is constructed by taking the linear span of all the vectors of the form $|\varphi^1\rangle \otimes |\psi^2\rangle$, where $|\varphi^1\rangle \in \mathcal{H}_1$ and $|\psi^2\rangle \in \mathcal{H}_2$. These tensor product vectors are prescribed in the following way: Let $|\varphi^1\rangle = (x_1, \ldots, x_d)^\top$, then

$$\left|\varphi^1\right\rangle \otimes \left|\psi^2\right\rangle = \left(x_1 \left|\psi^2\right\rangle, \ldots, x_d \left|\psi^2\right\rangle\right)^\top . \tag{2.3}$$

The inner-product on the composed space can be induced from the composing spaces through the following prescription,

$$\left\langle \varphi_x^1 \otimes \psi_x^2 \middle| \varphi_y^1 \otimes \psi_y^2 \right\rangle_{\mathcal{H}_1 \otimes \mathcal{H}_2} = \left\langle \varphi_x^1 \middle| \varphi_y^1 \right\rangle_{\mathcal{H}_1} \left\langle \psi_x^2 \middle| \psi_y^2 \right\rangle_{\mathcal{H}_2} . \tag{2.4}$$

We can also note that the operator space of the tensored Hilbert space $\mathcal{H}_1 \otimes \mathcal{H}_2$, $\mathcal{L}(\mathcal{H}_1 \otimes \mathcal{H}_2)$ is isomorphic to $\mathcal{L}(\mathcal{H}_1) \otimes \mathcal{L}(\mathcal{H}_2)$. Now, given two operators $X \in \mathcal{L}(\mathcal{H}, \mathcal{K})$ and $Y \in \mathcal{L}(\mathcal{V}, \mathcal{W})$, then the tensored operator $X \otimes Y \in \mathcal{L}(\mathcal{H} \otimes \mathcal{V}, \mathcal{K} \otimes \mathcal{W})$ is uniquely defined through its action on an arbitrary tensored state as $(X \otimes Y)|a\rangle \otimes |b\rangle = X|a\rangle \otimes Y|b\rangle$. Moreover, if $\{x_i\}$ are the eigenvalues of X and $\{y_j\}$ of Y, the eigenvalues of $X \otimes Y$ is given by $\{x_i y_j\}$.

Now, we can establish some shorthand notations regarding tensored objects. Tensored vectors $|x\rangle \otimes |y\rangle$ can be represented as $|x\rangle |y\rangle$, omitting the tensoring operation. It can alse be represented as $|xy\rangle$ when it is clear from the context and no confusion arises. When we have tensored spaces which are isomorphic to each other, that is $\mathcal{H}_1 = \mathcal{H}_2 = \mathcal{H}$, instead of writing $\mathcal{H} \otimes \mathcal{H}$ we can write $\mathcal{H}^{\otimes 2}$. This convention can also be used for identical operators on composing isomorphic spaces. For example $X \otimes X$ is written as $X^{\otimes 2}$.

Linear maps transforming operators

We delayed this discussion so as to introduce the notion of completely positive maps, which require tensor product structures for its definition. Given two Hilbert spaces, $\mathcal{H}$ and $\mathcal{K}$, we can have linear maps transforming operators on $\mathcal{H}$ to operators on $\mathcal{K}$. The set of such linear maps $\mathcal{M} : \mathcal{L}(\mathcal{H}) \to \mathcal{L}(\mathcal{K})$ is denoted as $\mathcal{L}(\mathcal{L}(\mathcal{H}), \mathcal{L}(\mathcal{K}))$. Given two maps $\mathcal{M} \in \mathcal{L}(\mathcal{L}(\mathcal{H}), \mathcal{L}(\mathcal{K}))$ and $\mathcal{N} \in \mathcal{L}(\mathcal{L}(\mathcal{V}), \mathcal{L}(\mathcal{W}))$, we can have the corresponding tensored map $\mathcal{M} \otimes \mathcal{N} \in \mathcal{L}(\mathcal{L}(\mathcal{H} \otimes \mathcal{V}), \mathcal{L}(\mathcal{K} \otimes \mathcal{W}))$, whose action is uniquely defined through its action on an arbitrary tensored operator as $(\mathcal{M} \otimes \mathcal{N})[X \otimes Y] = \mathcal{M}(X) \otimes \mathcal{N}(Y)$. When we have the same map $\mathcal{M}$ acting on composing isomorphic spaces, we can write it as $\mathcal{M}^{\otimes 2}$ instead of $\mathcal{M} \otimes \mathcal{M}$. Now, we can look at some relevant linear maps.

- ***Identity maps.*** The identity map, which is equivalent to the identity operator $\mathbf{id}$ in its action, is represented by $\mathcal{I}_{\mathcal{H}}$. It is a map $\mathcal{I}_{\mathcal{H}} \in \mathcal{L}(\mathcal{L}(\mathcal{H}))$ with the action $\mathcal{I}_{\mathcal{H}} : X \mapsto X$. When the space is clear from the context, we omit the subscript spaces on the identity map.

- ***Hermicity-preserving maps.*** A linear map $\mathcal{N} : \mathcal{L}(\mathcal{H}) \to \mathcal{L}(\mathcal{K})$ is said to be "Hermicity-preserving" if $\mathcal{N}(X)$ is self-adjoint whenever X is self-adjoint.

- ***Trace-preserving maps.*** A linear map $\mathcal{N} : \mathcal{L}(\mathcal{H}) \to \mathcal{L}(\mathcal{K})$ is said to be "trace-preserving" if $\mathrm{Tr}\{\mathcal{N}(X)\} = \mathrm{Tr}\{X\}$ for all $X \in \mathcal{L}(\mathcal{H})$.

- ***Unital maps.*** A linear map $\mathcal{N} : \mathcal{L}(\mathcal{H}) \to \mathcal{L}(\mathcal{K})$ is said to be "unital" if it maps the identity operator on $\mathcal{H}$ to the identity operator on $\mathcal{K}$, that is, $\mathcal{N}(\mathbf{id}_{\mathcal{H}}) = \mathbf{id}_{\mathcal{K}}$. We can note that trace-preserving unital maps exist only when the input and output spaces are identical.

- ***Positive maps.*** A linear map $\mathcal{N} : \mathcal{L}(\mathcal{H}) \to \mathcal{L}(\mathcal{K})$ is said to be positive when it maps each and every positive operator on $\mathcal{H}$ to a positive operator on $\mathcal{K}$, that is, $\mathcal{N}(\mathcal{L}_+(\mathcal{H})) \subseteq \mathcal{L}_+(\mathcal{K})$.

- ***Completely-positive maps.*** A linear map $\mathcal{N} : \mathcal{L}(\mathcal{H}) \to \mathcal{L}(\mathcal{K})$ is said to be "completely positive" when $\mathcal{I} \otimes \mathcal{N}(\mathcal{L}_+(\mathcal{H}_d \otimes \mathcal{H})) \subseteq \mathcal{L}_+(\mathcal{H}_d \otimes \mathcal{K})$ for all finite extensions $\mathcal{H}_d$.

A notion similar to the adjointness of operators is the adjointness of maps. Given a map $\mathcal{E}$, its adjoint or dual map $\mathcal{E}^\dagger$ is identified through the satisfaction of the equality $\mathrm{Tr}\{X\mathcal{E}(Y)\} = \mathrm{Tr}\left\{\mathcal{E}^\dagger(X)Y\right\}$ for all operators Y.

2.1.2 Choi–Jamiołkowski isomorphism

Choi–Jamiołkowski isomorphism [Cho75, Jam72] is a fundamental tool in this thesis, as it furnishes a unique representation of quantum processes and the higher-order quantum theoretic framework was developed based on this specific representation. Let $\mathcal{E} \in \mathcal{L}(\mathcal{L}(\mathcal{H}), \mathcal{L}(\mathcal{K}))$. Once we fix an orthonormal basis $\{|\varphi_k\rangle\}$ for the space d-dimensional space $\mathcal{H}$, we can have the bipartite vector $|\psi_+\rangle = \sum_{k=1}^{d} |\varphi_k\rangle |\varphi_k\rangle$. With this we can define the operator Ψ_+ as

$$\Psi_+ := |\psi_+\rangle\langle\psi_+| = \sum_{i,j} |\varphi_i\rangle\langle\varphi_j| \otimes |\varphi_i\rangle\langle\varphi_j| \, . \tag{2.5}$$

Then, we can have the map ChoiJ with the following action,

$$\mathsf{ChoiJ} : \mathcal{E} \mapsto \Phi_{\mathcal{E}} = (\mathcal{I} \otimes \mathcal{E})[\Psi_+]. \tag{2.6}$$

This map establishes the Choi-Jamiołkowski isomorphism between the space of linear maps $\mathcal{L}(\mathcal{L}(\mathcal{H}), \mathcal{L}(\mathcal{K}))$ and the space of linear operators $\mathcal{L}(\mathcal{H} \otimes \mathcal{K})$. Since we claim that ChoiJ furnishes the isomorphism we will have to present the inverse map to ChoiJ. The inverse map $\mathsf{ChoiJ}^{-1} : \Phi \mapsto \mathcal{E}_\Phi$ prescribes the action of the map $\mathcal{E}_\Phi$ as the following,

$$\mathcal{E}_\Phi[X] = \mathrm{Tr}_{\mathcal{H}} \left\{ \Phi_{\mathcal{E}} (X^\top \otimes \mathbf{id}_{\mathcal{K}}) \right\} . \tag{2.7}$$

A crucial aspect of the isomorphism is that completely positive maps are mapped to positive operators and vice versa.

2.1.3 Semi-definite programming

In a semi-definite programming task or semi-definite program (SDP), our task is to optimise a real-valued function on self-adjoint operators, constrained by linear functions [SC23]. An semi-definite program can be identified with a triple of objects: $\left(\mathsf{obj}, \{\Gamma_i^E\}, \{\Gamma_j^I\}\right)$, where obj, referred to as the “objective function”, is a real linear function which can always be written as $\mathsf{obj}(\cdot) = \mathrm{Tr}\{A \cdot\}$ for some self-adjoint A, $\{\Gamma_i^E\}$ is a finite collection of Hermicity-preserving maps prescribing linear equality constraints and $\{\Gamma_i^I\}$ is a finite collection of Hermicity-preserving maps prescribing linear inequality constraints. Given such a triple, what we want to carry out is the following procedure.

$$\begin{aligned} \text{maximise} \quad & \mathsf{obj}(X) = \mathrm{Tr}\{AX\} \, , & (2.8)\\ \text{subject to} \quad & \Gamma_i^E(X) = B_i \ \forall i \text{ and,} & (2.9)\\ & \Gamma_j^I(X) \leq C_j \ \forall j. & (2.10) \end{aligned}$$

This form of the program is referred to as of the "primal" form or the primal SDP. The set of primal variables X which satisfy Equation 2.9 and Equation 2.10 is referred to as the feasible set $\mathcal{F}$ associated with the SDP and the elements of this set are called feasible operators; If there does not exist a non-empty feasible set then there can be no solution to the SDP. As such, we can infer from the primal SDP that we are maximising the objective function over the feasible set. Those feasible operators X^* which maximise the objective function are referred to as the optimal operators. We can note that the optimal value $\alpha^* = \mathrm{Tr}\{AX^*\}$ is always lower bounded by $\alpha_j = \mathrm{Tr}\{AX'\}$ where X' are non-optimal operators. Here, we have presented the primal SDP as a maximisation task but we can equivalently cast it as a corresponding minimisation task by appropriately tinkering with the functions. Now, when the feasible set associated with an SDP is non-empty, we can always have an SDP which we call as the "feasibility" SDP. A feasibility SDP corresponding to the SDP we presented above is given by,

$$\begin{aligned} \text{find} \quad & X, && (2.11)\\ \text{subject to} \quad & \Gamma_i^E(X) = B_i \ \forall i \text{ and}, && (2.12)\\ & \Gamma_j^I(X) \leq C_j \ \forall j. && (2.13) \end{aligned}$$

Now, corresponding to every primal SDP, we can write down the "dual" SDP associated with it. For this recasting, we introduce dual variables Y_i and Z_j corresponding to each of the constraints $\Gamma_i^E(X) = B_i$ and $\Gamma_j^I(X) \leq C_j$. This enables us to write down the Lagrangian $\mathcal{L}$ which we associate to the SDP,

$$\begin{aligned} \mathcal{L} &= \mathrm{Tr}\{AX\} + \sum_i \mathrm{Tr}\left\{Y_i(B_i - \Gamma_i^E(X))\right\} + \sum_j \mathrm{Tr}\left\{Z_j(C_j - \Gamma_j^I(X))\right\} && (2.14)\\ &= \mathrm{Tr}\left\{X\left[A - \sum_i \Gamma_i^{E\dagger}(Y_i) - \sum_j \Gamma_j^{I\dagger}(Z_j)\right]\right\} + \\ &\qquad \sum_i \mathrm{Tr}\{Y_i B_i\} + \sum_j \mathrm{Tr}\{Z_i C_i\}. && (2.15) \end{aligned}$$

A natural constraint arises for Z_j with $Z_j \geq O$. Another constraint that is imposed is

$$A - \sum_i \Gamma_i^{E\dagger}(Y_i) - \sum_j \Gamma_j^{I\dagger}(Z_j) = O. \tag{2.16}$$

The Lagrangian for any feasible primal variable X which evaluated under the above constraints is given by $\mathcal{L} = \mathrm{Tr}\{Y_i B_i\} + \sum_j \mathrm{Tr}\{Z_i C_i\}$. Then, the dual SDP reads,

$$\text{minimise} \qquad \mathcal{L} = \mathrm{Tr}\left\{Y_i B_i\right\} + \sum_j \mathrm{Tr}\left\{Z_i C_i\right\}, \tag{2.17}$$

$$\text{subject to} \qquad A - \sum_i \Gamma_i^{E\dagger}(Y_i) - \sum_j \Gamma_j^{I\dagger}(Z_j) = O \text{ and,} \tag{2.18}$$

$$Z_j \geq O \ \forall j. \tag{2.19}$$

Here, the dual feasible set contains the tuples $(Y_1, Y_2, \ldots, Z_1, Z_2, \ldots)$ which simultaneously satisfy Equation 2.18 and Equation 2.19.

Weak and Strong duality

Let α^* be the optimal value to a given primal SDP and β^* be the one corresponding to the associated dual SDP. Then, "weak duality" dictates the satisfcation of $\alpha^* \leq \beta^*$. When "strong duality" is said to be held, then $\alpha^* = \beta^*$.

2.2 Building blocks of quantum theory

A quantum system describes some quantum degree of freedom of a physical system, and usually both of them are used synonymously. In the standard framework of quantum theory, each such quantum system $\mathcal{Q}$ is identified with an appropriate complex Hilbert space $\mathcal{H}_\mathcal{Q}$; the system and its Hilbert space are used interchangeably, thus the subscipt $\mathcal{Q}$ is dropped from now ownwards while referring to the system as just $\mathcal{H}$. The aforementioned appropriateness is prescribed by the nature of the degree of freedom as the dimension of the corresponding space. Moreover, if $\mathcal{H}_1$ and $\mathcal{H}_2$ are the corresponding Hilbert spaces of two quantum systems $\mathcal{Q}_1$ and $\mathcal{Q}_2$, then the Hilbert space of the composed system is identified with the tensor product Hilbert space $\mathcal{H}_1 \otimes \mathcal{H}_2$.

Now, quantum theory is a statistical or probabilistic physical theory, as in the predictions of the theory are probability distributions those describe the statistics observed in experiments. As such, the only comment quantum theory can make regarding a single recording of a particular measurement event is the probability with which that recording could occur. Regardless, any such probabilistic theory can be identified with two sets of mathematical objects, one containing states of the system and the other of measurement events, referred together as the "statistical duality" [BLPY16] and a pairing rule, combining one element from the first set and one from the second to furnish the probability of occurence of some possible physical event. In this section, we discuss the statistical duality of quantum theory: the set of states $\mathbf{St}(\mathcal{H})$ and the set of effects $\mathbf{Eff}(\mathcal{H})$ and associated combining rule, referred to as the Born rule; As we shall see, the states and effects of a particular system $\mathcal{H}$ are identified with specific operators on $\mathcal{H}$ and the rule prescribed through the trace functional on $\mathcal{L}(\mathcal{H})$. Moreover, we shall discuss the notions of quantum observables and processes.

2.2.1 Generating quantum probabilities: States and effects

Normalised vectors of $\mathcal{H}$ correspond to different physical preparations of the system, admitted by the theory. Among such vectors, those which differ to each other by a factor of phase $e^{i\theta}$ are indistinguishable from the predictions of the theory and as well as from experiments. Such an equivalent class of normalised vectors, corresponding to equivalent physical preparations, is identified as a single state of the system. The state of the system is then represented by a representative of that class $[|\psi\rangle] = \{|\psi\rangle, e^{i\theta}|\psi\rangle\}_\theta$, as $|\psi\rangle$. But this description of quantum states is restrictive as in that it fails to describe states which can be prepared by physically mixing of such states. The descriptor of states, that shall be discussed in the following, is general enough to rectify this.

States

Quantum states are assumed to contain all the information of a quantum system that can be extracted through quantum experiments. These states of a system $\mathcal{H}$ are identified with positive semi-definite, unit-trace linear operators on $\mathcal{H}$. Satisfying the desideratum that physical mixtures of states to be valid states, by construction the set of all states of the system $\mathcal{H}$ is convex and is identified with,

$$\mathbf{St}(\mathcal{H}) := \{\xi \in \mathcal{L}_+(\mathcal{H}) \text{ and } \mathrm{Tr}\{\xi\} = 1\}. \tag{2.20}$$

Unit-trace condition is often referred to as the "normalisation" of the state. Previously, the states we identified with vectors form a class referred to as "pure" states. Given any state vector $|\varphi\rangle$, we can associate it with its corresponding pure state operator $\varphi = |\varphi\rangle\langle\varphi|$; We can also observe that these pure states are projectors. These pure states are the extremal elements of the set and any other state can be written down as a non-unique convex mixture of pure states; These states are referred to as "mixed" states.

Effects

Given a quantum system prepared in a state ξ, as already discussed, quantum theory attempts to describe or characterise the measurement events associated with any quantum experiment or measurement procedure. An effect is a notion which is associated to each of any such measurement events and one which duly assigns a probability of occurence or recording to the associated event. In this sense, they can be identified with affine functionals on the set of states, $e : \mathbf{St}(\mathcal{H}) \to [0,1] \subset \mathbb{R}$. The action of this map is given by $e : \xi \mapsto \mathrm{Tr}\left\{E_e \xi\right\}$ where E_e is a positive operator bounded above by $\mathbf{id}$, which is associated to the map [HZ11]. With this identification, we can refer to these operators as effects. As such, we can define the set of effects on the system $\mathcal{H}$ as,

$$\mathbf{Eff}(\mathcal{H}) := \{E \in \mathcal{L}_+(\mathcal{H}) : E \leq \mathbf{id}_{\mathcal{H}}\}. \tag{2.21}$$

Quantum probabilities

As we have already discussed, a typical event in a quantum experiment reads like this: Given the system is prepared in a state ρ, you implement a measurement, on the system, that includes a measurement event or outcome m which is associated to the effect M. Then, quantum theory predicts the probability with which this measurement event m happens, as the conditional probability $p(m|\rho)$. This probability is given by the so-called Born rule as,

$$p(m|\rho) = \mathcal{B}(\rho, M) = \mathrm{Tr}\left\{\rho M\right\}. \tag{2.22}$$

As such the Born rule can be seen as a map $\mathcal{B} : \mathbf{St}(\mathcal{H}) \times \mathbf{Eff}(\mathcal{H}) \to [0,1] \subset \mathbb{R}$.

Entangled states and auxiliary systems

We have discussed that composite or multipartite quantum systems are identified with tensor product Hilbert space of the individual composing systems' Hilbert spaces. This tensorial structure of the space leads to the quantum feature called entanglement and its ramifications. For our purposes and within the scope of this thesis, it is adequate to identify and characterise entanglement for bipartite systems $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$. If a state vector $|\psi\rangle \in \mathcal{H}_1 \otimes \mathcal{H}_2$ can be written as $|\psi\rangle = |\psi_1\rangle \otimes |\psi_2\rangle$, for some $|\psi_1\rangle \in \mathcal{H}_1$ and $|\psi_2\rangle \in \mathcal{H}_2$, it is referred to as being "separable". State vectors belonging to this space, those are not separable are referred to as being "entangled". For example, consider the system $\mathbb{C}^2 \otimes \mathbb{C}^2$; the state $|\varphi\rangle \otimes |\psi\rangle$ is separable as well as the state $\frac{1}{\sqrt{2}}(|\varphi\varphi\rangle + |\varphi\psi\rangle)$, since it can be factored as $\frac{1}{\sqrt{2}}(|\varphi\rangle \otimes (|\varphi\rangle + |\psi\rangle))$. Not only we have entangled state vectors, we also have entangled operators on these tensored spaces. For us, the relevant operators are entangled states. Following [HZ11], a quantum state $\rho \in \mathbf{St}(\mathcal{H}_1 \otimes \mathcal{H}_2)$ is said to be "factored" if it can be expressed as $\rho = \rho_1 \otimes \rho_2$ for some $\rho_1 \in \mathbf{St}(\mathcal{H}_1)$ and $\rho_2 \in \mathbf{St}(\mathcal{H}_2)$. Moreover a state is said to be separable if it can be expressed as a convex sum of factored states. As such, we identify a state to be entangled if it is not separable. Now, we can identify the following relevant family of bipartite entangled states,

– ***Maximally entangled states.*** Given an orthonormal basis $\{|\varphi_k\rangle\}_{k=1}^{d}$, maximally entangled states on the bipartite system $\mathbb{C}^d \otimes \mathbb{C}^d$ is given by the family of pure states,

$$\frac{1}{\sqrt{d}} \sum_{j=1}^{d} |\varphi_j\rangle \otimes |\varphi_j\rangle \, . \tag{2.23}$$

As discussed in the introductory chapter of this thesis, entanglement has been found to be a resource in a plethora of quantum information processing tasks. In these tasks, entanglement is introduced by coupling the system of interest with another system, almost always of the same dimension, and preparing the composite system in entangled states. These extra systems which are coupled are referred to as "auxiliary" systems since they "assist" the tasks with their part in entanglement.

2.2.2 Measurement devices and Observables

A quantum measurement[3] is performed on a system to acquire some information about it and the physical setup implementing or realising such a measurement process is referred to as the associated measurement device. Analysing the notion, we can associate each and every quantum measurement with a collection of physical events which form the possible outcomes of the experiment. As such, we can associate a set of outcomes Ω to every measurement whose outcomes are well-characterised enough. In this thesis, outcome sets are restricted to be finite and discrete, without loss of generality, being subsets of the set of natural numbers $\mathbb{N}$. In general, these quantum measurements, or equivalently treated measurement devices, can be modeled by three equivalent [4] mathematical descriptors: as quantum observables, through quantum instruments (where post-measurement states are assumed to be available) or through quantum measurement models (where internal descriptions of the device are furnished when post-measurement states are assumed) [HZ11]. Within the scope of our thesis, we restrict quantum measurements to be described by observables.

[3]Considering the research that has been done in quantum information theoretic communities, the term "quantum measurement" might serve as an umbrella term for measurement or tomography of any quantum object, not necesseraliy of quantum states. Regardless, in any of such contexts, one should essentially measure quantum states to learn about the objects involved.

[4]With regard to the statistical aspect.

Observables

Historically, through out the development of quantum theory and afterwards, the word "observable" has been used to refer to measurable physical quantities, for example like energy and position and were represented by self-adjoint operators [Per97] and measuring these observables on a system corresponded with measurements of these quantitites. Later on, this notion of an observable was generalised to have operational interpretations of the experiment [DL70] as discussed above. Subsequently, these generalised notion of observables were identified with specific "normalised positive operator-valued measures" (POVMs) or, since we have introduced the notion of effects, "normalised effect-valued measures". As a consequence, these generalised quantum observables describe "observable" statistics of any quantum experiment. As such, in this thesis, we refer to these normalised effect-valued measures as observables. Now, the set of all observables on a system $\mathcal{H}$, with an outcome set Ω is defined as,

$$\mathbf{Obs}(\Omega, \mathcal{H}) := \left\{ \mathsf{M} : \Omega \to \mathbf{Eff}(\mathcal{H}) \ ; \ \sum_{\omega_k \in \Omega} \mathsf{M}(\omega_k) = \mathbf{id}_{\mathcal{H}} \right\}. \tag{2.24}$$

Some classes of observables

We define and discuss certain classes of observables, those are relevant as well as appear in the investigations of this thesis. As a shorthand notation, we denote $\underline{n}$ as the set $\{1, \ldots, n\} \subset \mathbb{N}$.

- ***Sharp and unsharp observables.*** An observable $\mathsf{P} \in \mathbf{Obs}(\underline{n}, \mathbb{C}^d)$ is said to be *sharp* or *projective* when each of the underlying effects is a projector, that is $\mathsf{P}(j)^2 = \mathsf{P}(j)$ for all $j \in \underline{n}$; Note that for sharp observables, $n \leq d$. On the other hand, an observable $\mathsf{Q} \in \mathbf{Obs}(\underline{n}, \mathbb{C}^d)$ is said to be *unsharp* or *non-projective* if it is not a sharp observable. That is, there should exist at least one $j \in \underline{n}$ such that $\mathsf{Q}(j)^2 \neq \mathsf{Q}(j)$.

- ***Rank-1 observables.*** An observable $\mathsf{M} \in \mathbf{Obs}(\underline{n}, \mathbb{C}^d)$ is said to be *rank-1* when each of the underlying effects is a rank-1 operator.

- ***von Neumann observables.*** An observable $\mathsf{M} \in \mathbf{Obs}(\underline{d}, \mathbb{C}^d)$ is said to be *von Neumann* if it is both sharp as well as rank-1. As such, for a given system, all von Neumann observables have the same number of outcomes, which equals the dimension of the system.

- ***Full-rank observables.*** An observable $\mathsf{M} \in \mathbf{Obs}(\underline{n}, \mathbb{C}^d)$ is said to be *full-rank* if $\mathsf{M}(k)$ is a full-rank operator, for all $k \in \underline{n}$.

- ***Rank-deficient observables.*** An observable is defined to be *rank-deficient* if atleast one of its effects is a rank-deficient operator.

- ***Binary, non-binary and trivial observables.*** An observable M is defined to be *binary* when $\mathsf{M} \in \mathbf{Obs}(\underline{2}, \mathbb{C}^d)$, that is, it is composed of two effects. On the other hand, any observable $\mathsf{N} \in \mathbf{Obs}(\underline{n}, \mathbb{C}^d)$ with $n \geq 3$ is referred to as *non-binary*. Trivial observables are those with just one effect, included in $\mathbf{Obs}(\underline{1}, \mathbb{C}^d)$. We can note that we cannot learn anything about the system by implementing any trivial observable since they just result in producing the probability 1 on any state.

- ***Coin-toss observables.*** An observable $\mathsf{M} \in \mathbf{Obs}(\underline{n}, \mathcal{H})$ is a coin toss observable when its effects take the following form: $\mathsf{M}(1) = q_1\mathbf{id}, \mathsf{M}(2) = q_2\mathbf{id}, \ldots, \mathsf{M}(n) = q_n\mathbf{id}$ where $(q_1, q_2, \ldots, q_n) = \vec{q}$ is some probability distribution. With such an observable we cannot learn anything about the system since the measurement will always result in a statistics described by $\vec{q}$, regardless of the state. A coin-toss observable is said to be unbiased if the associated $\vec{q}$ is the flat distribution $(\frac{1}{n}, \frac{1}{n}, \ldots, \frac{1}{n})$ and biased if otherwise.

- ***Informationally complete observables.*** An observable $\mathsf{M} \in \mathbf{Obs}(\underline{n}, \mathbb{C}^d)$ is defined to be informationally complete if the linear span of its effects cover the whole operator space, i.e., $\boldsymbol{span}\{\mathsf{M}(1), \ldots \mathsf{M}(n)\} = \mathcal{L}(\mathbb{C}^d)$. This essentialy means that any state $\rho \in \mathbf{St}(\mathbb{C}^d)$ can be tomographed completely from the statistics produced by such an observable, when it has been used to measure the state.

Postprocesing of observables

Classical postprocessing [MdM90, Hei05] is a procedure by which we can obtain an observable from another by coarse-graining the later in the following sense. We can say that an observable $\mathsf{B} \in \mathbf{Obs}(\Omega_\mathsf{B}, \mathcal{H})$ is a post-processing of another observable $\mathsf{A} \in \mathbf{Obs}(\Omega_\mathsf{A}, \mathcal{H})$ if there exists a column stochastic matrix μ_{xy} such that

$$\mathsf{B}(x) = \sum_{y \in \Omega_\mathsf{A}} \mu_{xy} \mathsf{A}(y). \tag{2.25}$$

The matrix μ is referred to as the associated postprocessing matrix. As such, postprocessings can be viewed as linear maps $\mathcal{P}_\mu : \mathbf{Obs}(\Omega_\mathsf{A}, \mathcal{H}) \to \mathbf{Obs}(\Omega_\mathsf{B}, \mathcal{H})$. We can note that such postprocessings are not the only way we can classically obtain an observable from other observables; When we have a finite collection of observables, we can mix them to arrive at new ones. More generally, we can postprocess a collection of observables and mix them; This procedure is referred to as "simulation" of that particular collection of observables.

Incompatibile observables

Incompatibility or compatibility of a finite collection of observables addresses the question whether there exists a joint measurement or observable corresponding to this collection. If there exists such a joint or "parent" observable, then from the statistics it furnishes, we can reconstruct the statistics of each of the observables belonging to the collection; And, this should hold for all possible experiments. If a collection admits a joint observable, then it is referred to as being jointly measurable or "compatible"; Otherwise, it is referred to as being incompatible. From the aforementioned statistical definition of incompatibility, we can note that two observables $\mathsf{A} \in \mathbf{Obs}(\Omega_{\mathsf{A}}, \mathcal{H})$ and $\mathsf{B} \in \mathbf{Obs}(\Omega_{\mathsf{B}}, \mathcal{H})$ are compatible if each of them can be separately postprocessed from a single joint observable. That is, if there exists an observable $\mathsf{C} \in \mathbf{Obs}(\Omega_{\mathsf{C}}, \mathcal{H})$ and postprocessings $\mathcal{P}_\mu$ and $\mathcal{P}_\nu$ such that $\mathcal{P}_\mu(\mathsf{C}) = \mathsf{A}$ and $\mathcal{P}_\nu(\mathsf{C}) = \mathsf{B}$ are satisfied. When this is satified, we can equivalently construct an observable $\mathsf{Q} \in \mathbf{Obs}(\Omega_{\mathsf{A}} \times \Omega_{\mathsf{B}}, \mathcal{H})$ such that we can get the observables through the following process of marginalisation [HMZ16],

$$\sum_{x \in \Omega_{\mathsf{A}}} \mathsf{Q}(x, y) = \mathsf{B}(y) \quad \text{and} \quad \sum_{y \in \Omega_{\mathsf{B}}} \mathsf{Q}(x, y) = \mathsf{A}(x). \tag{2.26}$$

This notion of incompatibility can be extended to anu finite collection of observables.

2.2.3 System evolutions: Quantum processes

As discussed earlier, since quantum theory is a probabilistic physical theory, general physical evolutions of a quantum system should be modeled to alter observed statistics. Since the statistics is prescribed by pairings of a state and an effect, as $\mathrm{Tr}\{\rho E\}$, we can model the mathematical descriptor of the physical evolution as either acting on the state or on the effects. This descriptor is referred to as quantum operations. Moreover, lossless quantum operations are referred to as quantum channels. In the rest of this thesis, our references to quantum channels can be replaced with quantum operations with appropriate context changes. As we were discussing, evolutions of the system can be modeled in two ways as described by the following "pictures" dual to each other.

– ***Schrödinger picture.*** When channels are modeled as to transform states on $\mathcal{H}$ to states on $\mathcal{K}$, they are identified with completely-positive (CP), trace-preserving (TP) linear maps, $\mathcal{E} : \mathcal{L}(\mathcal{H}) \to \mathcal{L}(\mathcal{K})$. Since we shall work majorly with channels in this picture, we define the set of all channels transforming states of a system $\mathcal{H}$ to states of a system $\mathcal{K}$ is defined as

$$\mathbf{Ch}(\mathcal{H}, \mathcal{K}) := \{\mathcal{E} \in \mathcal{L}(\mathcal{L}(\mathcal{H}), \mathcal{L}(\mathcal{K})) \ : \ \mathcal{E} \text{ is TP and CP}\}. \tag{2.27}$$

In this light, quantum operations are completely-positive trace-non-increasing maps.

– ***Heisenberg picture.*** On the other hand, equivalently, when they are modeled as to transform effects, they transform effects on $\mathcal{K}$ to effects on $\mathcal{H}$. These are identified with completely-positive unital linear maps. Now, given a channel $\mathcal{E}$ in the Schrödinger picture, its corresponding dual map in the Heisenberg picture is represented as $\mathcal{E}^*$ which coincides with the adjoint map of $\mathcal{E}$.

Traditionally, quantum operations and consequently quantum channels were used to refer to quantum evolutions and concatenated evolutions where none of the evolutions were separately correlated to each other through quantum memory systems. These later generalised descriptions of quantum evolutions were initially modeled in [HN03, KW05] and were referred to as quantum channels with memory [KW05]. They are themselves quantum operations satisfying corresponding defining conditions only to satisfy extra conitions those reflect their internal causal structures[5]. In the rest of the thesis, when we refer to "quantum processes", we essentially include these generalised descriptions as well.

Representations of channels

– ***As an operator-sum.*** Given a channel $\mathcal{E} \in \mathbf{Ch}(\mathcal{H}, \mathcal{K})$, there exists at least one finite collection of non-zero operators $\{S_1, \ldots, S_n\} \subset \mathcal{L}(\mathcal{H} \otimes \mathcal{K})$ satisfying $\sum_j S^\dagger S = \mathbf{id}_h$, such that the action of the channel is prescribed by this collection as,

$$\mathcal{E}(\xi) = S_1 \xi S_1^\dagger + \cdots + S_n \xi S_n^\dagger. \tag{2.28}$$

– ***Due to Choi–Jamiołkowski isomorphism .*** Given a channel $\mathcal{E} \in \mathbf{Ch}(\mathcal{H}, \mathcal{K})$, the Choi–Jamiołkowski isomorphism prescribes the operator $\Phi_\mathcal{E} \in \mathcal{L}(\mathcal{H} \otimes \mathcal{K})$ as follows. Let $\{|k\rangle\}$ be an orthonormal basis for $\mathcal{H}$, then $\tilde{\Psi}_+ = |\psi_+\rangle\langle\psi_+|$ where $|\psi_+\rangle = \sum_k |k\rangle \otimes |k\rangle$ is the maximally entangled non-normalised state vector. Then, due to the Choi–Jamiołkowski isomorphism , the *Choi-Jamiołkowski operator*[6] $\Phi_\mathcal{E}$ of $\mathcal{E}$ is given by,

$$\Phi_\mathcal{E} = \mathcal{I} \otimes \mathcal{E}(\Psi_+). \tag{2.29}$$

[5]This notion of memory channels will be discussed later on in this chapter.

[6]When the normalised maximally entangled state is used in the prescription, the corresponding operator is referred to as the *Choi state* of the channel.

Complete positivity of $\mathcal{E}$ is translated as the positivity of the operator $\Phi_{\mathcal{E}}$ whereas trace-preservation takes the form,

$$\mathrm{Tr}_{\mathcal{K}}\{\Phi_{\mathcal{E}}\} = \mathbf{id}_{\mathcal{H}}. \tag{2.30}$$

Now, given the Choi-Jamiołkowski operator $\Phi_{\mathcal{E}}$, the corresponding channel, or equivalently the action of the channel, is retrieved through the inverse Choi-Jamiołkowski isomorphism as

$$\mathcal{E}(X) = \mathrm{Tr}_{\mathcal{H}}\{\Phi_{\mathcal{E}}(X^{\top} \otimes \mathbf{id}_{\mathcal{K}})\}. \tag{2.31}$$

Combining channels

Given two channels $\mathcal{E} \in \mathbf{Ch}(\mathcal{H}_1, \mathcal{H}_2)$ and $\mathcal{F} \in \mathbf{Ch}(\mathcal{H}_3, \mathcal{H}_4)$, we can combine them in two possible ways to arrive at a resultant channel.

– ***Parallel combination.*** The first one is to combine them in parallel, resulting in the tensored channel $\mathcal{G} = \mathcal{E} \otimes \mathcal{F} \in \mathbf{Ch}(\mathcal{H}_1 \otimes \mathcal{H}_3, \mathcal{H}_2 \otimes \mathcal{H}_4)$. To elaborate, this corresponds to individually implementing $\mathcal{E}$ and $\mathcal{F}$ locally on the bipartite system $\mathcal{H}_1 \otimes \mathcal{H}_3$. As such, any two channels, irrespective of their corresponding input and output spaces, can be combined in parallel. The Choi-Jamiołkowski operator of the parallely combined channel, $\Phi_{\mathcal{G}}$, is given by the tensored Choi-Jamiołkowski operators of the composing channels as,

$$\Phi_{\mathcal{G}} = \Phi_{\mathcal{E}} \otimes \Phi_{\mathcal{F}}. \tag{2.32}$$

– ***Sequential combination.*** Sequential combination corresponds to the implemention of one channel $\mathcal{F}$ right after implemention of the other $\mathcal{E}$, represented as $\mathcal{F} \circ \mathcal{E} \in \mathbf{Ch}(\mathcal{H}_1, \mathcal{H}_4)$. Consequently, unlike parallel combination, for sequentially combining two channels, the second one's input space should be identical to the first ones's output space. That is, for $\mathcal{F} \circ \mathcal{E}$ is admissible only when $\mathcal{H}_2 = \mathcal{H}_3$. Moreover, when all the involved input and output spaces are identical, we can sequentially combine the two channels in two possible orders, $\mathcal{F} \circ \mathcal{E}$ and $\mathcal{E} \circ \mathcal{F}$. The Choi-Jamiołkowski operator of the sequentially combined channel $\Phi_{\mathcal{G}} = \mathcal{F} \circ \mathcal{E}$ is given by the so called "link product" [CDP08c] between the composing Choi operators $\Phi_{\mathcal{F}} * \Phi_{\mathcal{E}}$ given by,

$$\Phi_{\mathcal{F}} * \Phi_{\mathcal{E}} = \mathrm{Tr}_{\mathcal{H}_2}\left\{(\Phi_{\mathcal{E}} \otimes \mathbf{id}_{\mathcal{H}_4})(\Phi_{\mathcal{F}} \otimes \mathbf{id}_{\mathcal{H}_1})\right\}. \tag{2.33}$$

Here, we can note that the operators, $(\Phi_{\mathcal{E}}\otimes\mathbf{id}_{\mathcal{H}_4})$ and $(\Phi_{\mathcal{F}}\otimes\mathbf{id}_{\mathcal{H}_1})$, are represented as if they belong to different tensored spaces. But these spaces are equivalent or isomorphic to each other with appropriate reordering of the component spaces. As such, the product $(\Phi_{\mathcal{E}} \otimes \mathbf{id}_{\mathcal{H}_4})(\Phi_{\mathcal{F}} \otimes \mathbf{id}_{\mathcal{H}_1})$ should be considered as legit with the consensus that the multiplication happens with appropriate reordering of the underlying spaces.

Some relevant channels

Here we discuss few of the channels which are relevant for this thesis.

- ***Unitary channels.*** Unitary channels describe reversible evolutions of a system. The action of a unitary channel $\mathcal{U}$ is dictated by an associated unitary operator U as,
$$\mathcal{U}(\rho) = U\rho U^\dagger. \tag{2.34}$$
- ***Measure-and-prepare channels.*** Measure-and-prepare channnels describe preparation of a system $\mathcal{K}$ conditioned on the outcomes of a measurement carried out on another system $\mathcal{H}$. Consequently, such a channel $\mathcal{N} \in \mathbf{Ch}(\mathcal{H},\mathcal{K})$ is characterised by an observable $\mathsf{N} \in \mathbf{Obs}(\underline{n},\mathcal{H})$ and a finite collection of n states $\{\xi_1,\ldots,\xi_n\} \subset \mathbf{St}(\mathcal{K})$. It is defined through the action on an arbitrary state $\rho \in \mathbf{St}(\mathcal{H})$,
$$\mathcal{N}(\rho) = \sum_{i=1}^{n} \mathrm{Tr}\{\mathsf{N}(i)\rho\}\xi_i. \tag{2.35}$$
- ***State preparation channels.*** For every state $\xi \in \mathbf{St}(\mathcal{H})$, there exists its corresponding preparation channel $\mathcal{P}_\xi \in \mathbf{Ch}(\mathbb{C},\mathcal{H})$, whose action is given by
$$\mathcal{P}_\xi : c \mapsto \xi. \tag{2.36}$$
Choi-Jamiołkowski operators of these channels are the associated states themselves.
- ***Measurement channels.*** For every observable $\mathsf{M} \in \mathbf{Obs}(\Omega,\mathcal{H})$, there exists its corresponding measurement channel $\Lambda_{\mathsf{M}} \in \mathbf{Ch}(\mathcal{H},\mathbb{C}^{|\Omega|})$, whose action, provided $\{|\varphi_k\rangle\}$ being an orthonormal basis for $\mathbb{C}^{|\Omega|}$, is given by
$$\Lambda_{\mathsf{M}} : \varrho \mapsto \sum_{k\in\Omega} \mathrm{Tr}[\varrho\mathsf{M}(k)]\,|\varphi_k\rangle\!\langle\varphi_k|\,. \tag{2.37}$$
We can notice that these channels form a special case of measure-and-prepare channels, where the collection of states which are conditionally prepared form an orthonormal basis. In fact, we can see that these "quantum-to-classical" channels are defined in such a way that they encode the classical probability distribution furnished by the observable M on a state ϱ

onto the state $\Lambda_{\mathsf{M}}(\varrho)$. By performing an observable composed of the effects $\{|\varphi_k\rangle\langle\varphi_k|\}$, we can observe the same statistics, essentially retreiving it back from the state.

Choi operator of measurement channels

We devote here space to evaluate the Choi-Jamiołkowski operator of measurement channels since they are the central objects we shall work with during our investigations in chapter 4 and chapter 5.

$$\begin{aligned}
\mathcal{I}\otimes\Lambda_{\mathsf{M}}(\Psi_+) &= \sum_{x,y} |x\rangle\langle y| \otimes \Lambda_{\mathsf{M}}(|x\rangle\langle y|) && (2.38)\\
&= \sum_{x,y} |x\rangle\langle y| \otimes \sum_k \mathrm{Tr}\left\{\mathsf{M}(k)\,|x\rangle\langle y|\right\} |k\rangle\langle k| && (2.39)\\
&= \sum_k \sum_{x,y} \langle y|\mathsf{M}(k)x\rangle\, |x\rangle\langle y| \otimes |k\rangle\langle k| && (2.40)\\
&= \sum_k \mathsf{M}(k)^\top \otimes |k\rangle\langle k|\,. && (2.41)
\end{aligned}$$

2.3 Higher-order quantum theory

In this section, we sketch the higher-order theoretic notions, within the scope of what all are relevant for this thesis; Specifically we discuss what are quantum combs and testers. For a much more detailed treatment of this subject, one can refer to [CDP09, Zim08, GW07].

Combining quantum building blocks

Prior to proceeding to our main discussions, we look at how we can "combine" the basic building blocks of quantum theory which we sketched previously. Subsequently, we establish conventions on how we illustrate these notions. Without loss of generality, we can consider the following scheme.

- A bipartite quantum state $\xi \in \mathbf{St}(\mathcal{H}\otimes\mathcal{J})$ is prepared. Such states or state preparations are illustrated with semi-ellipses with their bases facing right.
- A quantum process $\mathcal{E} \in \mathbf{Ch}(\mathcal{H},\mathcal{K})$ is implemented on the system $\mathcal{H}$, transforming the state to $\mathcal{I}\otimes\mathcal{E}(\xi)$. Such processes are illustrated with rounded-edge boxes.
- The state is then measured using an n-outcome observable $\mathsf{M} \in \mathbf{Obs}(\underline{n},\mathcal{K}\otimes\mathcal{J})$. Such observables are illustrated with semi-ellipses with their bases facing left.

– The measurement process results in the statistics, which is described by the discrete conditional probability distribution $p(k|\mathsf{M}, \mathcal{E}, \xi) = \mathrm{Tr}\left\{(\mathcal{I} \otimes \mathcal{E}(\xi))\mathsf{M}(k)\right\}$. This classical information is illustrated using double lines.

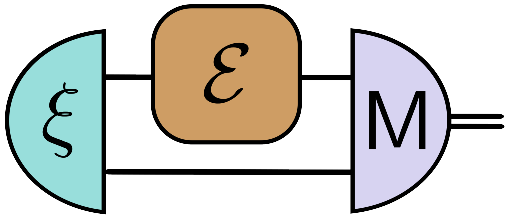

Figure 2.1: Prepared quantum state ξ are illustrated as semi-ellipses with their base facing right. Single-stroked lines correspond to quantum systems. In this thesis, time flows left to right in illustrations. The two lines coming out of the state semi-ellipse tells us that it is composed of two systems. A quantum channel $\mathcal{E}$ is implemented on of the systems; Shown as the system line entering and exiting the rounded-edge box representing $\mathcal{E}$. An observable M is represented using semi-ellipse with base facing left. This illustrates a measurement performed on the bipartite system (the two lines going inside). The closely spaced double lines, coming out, correspond to the classical information we record from the measurement.

Creating higher orders

As we have already discussed, a particular quantum state of a system is assumed to possess all the informational content that can be learned about the system, retrievable through different possible measurements. Moreover, evolutions of the system can also be conceived as state transformations described by quantum channels. In this choice of context, we can see that states are being considered as the central objects of interest, whose information content we want to extract through measurements and those undergo transformations. In fact, there are instances where we would want to ask questions regarding quantum channels so as to learn from them. There are also instances those can be considered and described as transformations of these channels. These are scenarios where, instead of quantum states, state transformations are considered as the central objects. This procedure of replacing the central objects with their transformations can be done to create hierarchial orders within quantum theory. Let us consider the studies concerning states, as central objects, as belonging to 0^{th} order quantum theory. Then, those with quantum channels as central objects as belonging to 1^{st} order quantum theory. Following the pattern, 2^{nd} order theory deals with transformations of quantum channels or superchannels [CDP08a] as the central objects. This procedure can be duly continued to identify and create higher orders, which is reflected in the following figure. This approach of acknowleding higher-orders is referred to as the "axiomatic approach" [CDP09].

2.3.1 Quantum combs

We discuss and give characterisations of general quantum causal processes, those quantum processes with definite causal ordering of subprocesses; We assert that these quantum processes are exactly the "central objects" we discussed in the previous passage. Prior to that, we can have a brief overview of quantum memory channels and the system indexing convention that shall be used for the rest of this thesis.

Memory channels and Pavia indexing

Kretschmann and Werner modeled quantum channels with memory in their work [KW05]. This modeling is adequately general to describe any quantum process whose outputs located in time till t are independent on all the inputs located in time greater than t; Consequently, these processes are composed of causally ordered sequence of discrete time steps. Any such process is referred to as being composed of memory channels, when there are "memory" quantum systems connecting such composing "single time-step channels"[7], as shown in Figure 2.2. Moreover, this notion of memory channels enables us to characterise them with the number of time steps they possess.

In this thesis, we adopt the following indexing convention which we refer to as the "Pavia indexing convention"[8]. As we can see in the illustrative example presented in Figure 2.2, given a memory channel of interest, we index input systems using even numbers, once we start with 0 and correspondingly the output systems with odd numbers. Moreover, the ordering of these natural numbers reflect the ordering of the composing channels in time and the associated memory channel structure of the process.

Transforming single time-step quantum processes

Single time-step quantum processes are nothing channels without any discernable internal " discrete time-step" structure. Since we had a discussion on the

[7] In general, these channels can be replaced by transformations which result in randomised outputs, described by the notion of "quantum instruments" [HZ11]. Within the scope of this thesis, all processes are assumed to be deterministic.

[8] Since it is found to be first used in [CDP08d] and adopted in subsequent works, those developed HOQT notions, by authors who hailed from the University of Pavia.

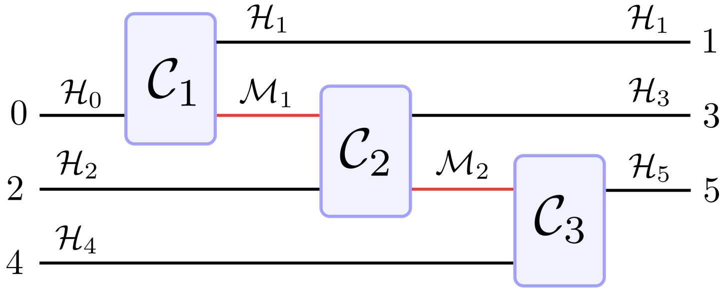

Figure 2.2: A concatenation of three memory channels $\mathcal{C}_1 \in \mathbf{Ch}(\mathcal{H}_0, \mathcal{H}_1 \otimes \mathcal{M}_1)$, $\mathcal{C}_2 \in \mathbf{Ch}(\mathcal{M}_1 \otimes \mathcal{H}_2, \mathcal{H}_3 \otimes \mathcal{M}_2)$ and $\mathcal{C}_3 \in \mathbf{Ch}(\mathcal{M}_2 \otimes \mathcal{H}_4, \mathcal{H}_5)$, where $\mathcal{M}_1$ and $\mathcal{M}_2$ are quantum memory systems (represented by red lines). In the Pavia indexing system, we denote the ingoing systems of the composite channel with even numbers, but starting with 0 and correspondingly the outgoing systems with odd numbers. This reflects the memory channel structure of the process as well; For example, the output system "1" exists in time before the input system "2".

perspective of creating higher-orders, we look the admissible transformations of these channels [CDP08a]; They have to transform channels to valid channels as well transform any local channels to valid local channels, similar to the conditions quantum channels themselves have to satisfy when it comes to describing state transformations. Regardless, we can have the following illustration for the most general transformation of a channel $\mathcal{C}$.

– The most general transformations are realised by a pair of post-processing or pre-processing $\mathcal{C}$ by $\mathcal{P}_1$ and $\mathcal{P}_2$, respectively. It is possible that there can be different pairs resulting in the same transformation. This enables us to consider such networks as equivalent ones, those result in identical transformations for identical inputs.

– Such a transformative network is illustrated in Figure 2.3. We note that such a network is, by itself, a memory channel with two time steps.

Quantum combs and their normalisation

We have seen that the most general transformation of a single time-step quantum process is mediated by a memory channel with two time-steps. We could continue this exercise, to identify the most general transformations those transform an N time-step memory channel to a single time-step one; We can extrapolate and find that those are mediated by $(N+1)$ time-step memory chennels. Now, every memory channel can always be viewed as an incomplete quantum network or circuit with open slots. This is illustrated in Figure 2.4. Given a memory channel, the Choi-Jamiołkowski operator of this channel or, equivalently, the incomplete

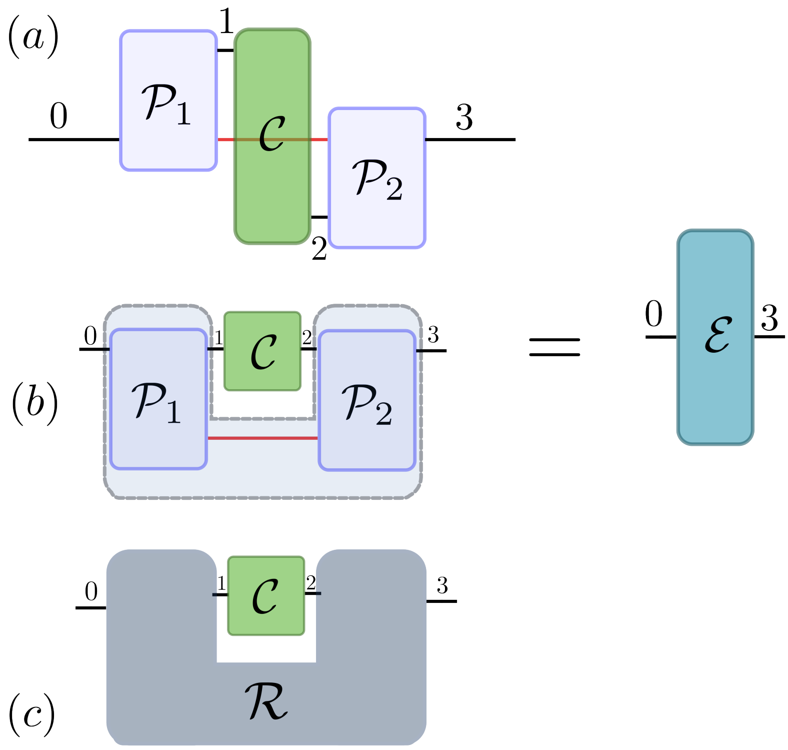

Figure 2.3: (a) shows the post general quantum network or circuit transforming a channel $\mathcal{C} \in \mathbf{Ch}(1,2)$ into another channel $\mathcal{E} \in \mathbf{Ch}(0,3)$, through sandwiching of appropriate channels $\mathcal{P}_1$ and $\mathcal{P}_2$; We note that this is a memory channel in general. This network can be redrawn into (b) so as to look like a "comb", where you plug-in $\mathcal{C}$ and get an output $\mathcal{E}$. The whole memory channel can be seen as a single quantum network $\mathcal{R}$, with a single open slot where you plug-in $\mathcal{C}$; In fact, $\mathcal{R}$ represents any memory channel that could transform $\mathcal{C}$ into an $\mathcal{E}$.

quantum network is referred to as the "quantum comb"[9] of this network. Moreover, two incomplete networks which results in identical transformations for all identical inputs are equivalent to each in this regard. That is, even if the component subprocesses are different, the Choi operator of the collective incomplete networks are identical to each other. As such, a quantum comb corresponds to all such equivalent networks. We can note that a quantum comb with (N) time-steps is characterised by $(N-1)$ open slots, into which you can "plug-in" appropriate quantum processes so as to transform them.

An $(N-1)$-slot quantum comb $\Phi_{\mathcal{R}}^{(N)}$ is essentially the Choi-Jamiołkowski operator of a quantum channel $\mathcal{R}^{(N)} \in \mathbf{Ch}(\mathcal{H}_0 \otimes \mathcal{H}_2 \otimes \cdots \otimes \mathcal{H}_{2N-2}, \mathcal{H}_1 \otimes \mathcal{H}_3 \otimes \cdots \otimes \mathcal{H}_{2N-1})$. On top of the constraints $\Phi_{\mathcal{R}}^{(N)}$ satisfy as a bona fide channel, it

[9] They look like hair combs.

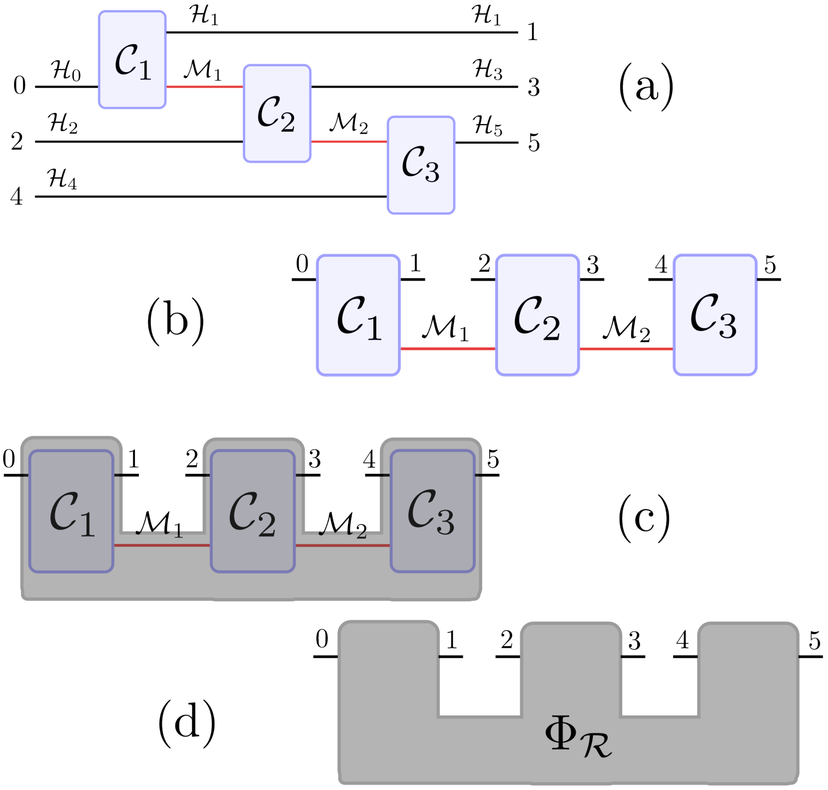

Figure 2.4: The memory channel in (a) is having three time steps. This memory channel can be visualised as (b) by appropriate redrawings respecting the causal structure. Moreover, this incomplete network can be visualised as a single object in (c), as having the form of a "comb" and characterised by the two open slots. In (d), we can see that the Choi-Jamiołkowski operator of the associated memory channel, $\Phi^{(2)}_{\mathcal{R}}$, is a single descriptor of it. This particular memory channel $\mathcal{R}$ is composed internally by single time step channels $\mathcal{C}_1, \mathcal{C}_2$ and $\mathcal{C}_3$. But, it is possible that some other triples of channels can also result in the same $\Phi^{(2)}_{\mathcal{R}}$. This is expressed in (d) where we are ignorant of its internal componets but care about the transformations it can result in, which depends soley on $\Phi^{(2)}_{\mathcal{R}}$.

also satisfies the following recursive "normalisation" conditions which reflect the causal structure of the network,

$$\mathrm{Tr}_{2k-1}\left\{\Phi^{(k)}_{\mathcal{R}}\right\} = \Phi^{(k-1)}_{\mathcal{R}} \otimes \mathbf{id}_{2k-2} \quad ; \quad k \in \{0, \dots, N\} \tag{2.42}$$

$$\mathrm{Tr}_1\{\Phi^{(1)}_{\mathcal{R}}\} = \mathbf{id}_0. \tag{2.43}$$

Here, the "normalisation objects" $\Phi^{(k)}_{\mathcal{R}}$ corresponds to combs of subnetworks $\mathcal{R}^{(k)}$. Now onwards, whenever we refer to a N-slot quantum comb, the sequence of input and output spaces and their associated dimensions are understood to be fixed within contexts. As such, corresponding set of combs will be convex in nature.

Linking combs and the generalised link product

We can "link" two combs by plugging-in one, which we require to be transformed, to another, which facilitates this transformation. To realise such valid link ups, both of the combs should require appropriate sequence of input and output systems. In such cases, we can define the following "generalised" link product between the combs. Let $\Phi_{\mathcal{R}} \in \mathcal{L}(\bigotimes_{j\in I_{\mathcal{R}}} \mathcal{H}_j)$ and $\Phi_{\mathcal{S}} \in \mathcal{L}(\bigotimes_{k\in I_{\mathcal{S}}} \mathcal{H}_k)$ be two combs on respective index sets $I_{\mathcal{R}}$ and $I_{\mathcal{S}}$. The generalised link product $\Phi_{\mathcal{R}} * \Phi_{\mathcal{S}}$ is then the following Choi-Jamiołkowski operator of the composed processes $\mathcal{S} \circ \mathcal{R}$, acts on the space $\mathcal{H}_{I_{\mathcal{S}}\setminus I_{\mathcal{R}}} \otimes \mathcal{H}_{I_{\mathcal{R}}\setminus I_{\mathcal{S}}}$, where we have adopted the notation $\mathcal{H}_I = \bigotimes_{x\in I} \mathcal{H}_x$,

$$\Phi_{\mathcal{S}} * \Phi_{\mathcal{S}} = \mathrm{Tr}_{I_{\mathcal{R}}\cap I_{\mathcal{S}}} \left\{ (\mathbf{id}_{I_{\mathcal{S}}\setminus I_{\mathcal{R}}} \otimes \Phi_{\mathcal{R}}^{\top_{I_{\mathcal{R}}\cap I_{\mathcal{R}}\mathcal{S}}})(\mathbf{id}_{I_{\mathcal{R}}\setminus I_{\mathcal{S}}} \otimes \Phi_{\mathcal{S}}) \right\}, \tag{2.44}$$

with $\top_{\mathcal{H}}$ being the transpose operation defined with a chosen basis of the space $\mathcal{H}$. A particular form of the generalised link product between two combs is when they cannot be linked as such, that is when $I_{\mathcal{R}} \cap I_{\mathcal{S}} = \varnothing$. In such cases, $\Phi_{\mathcal{S}} * \Phi_{\mathcal{R}} = \Phi_{\mathcal{S}} \otimes \Phi_{\mathcal{R}}$. As we shall see, we study exclusively combs of these forms in chapter 4 and chapter 5. We note that, since these combs cannot be linked together through any system, such combs are "memoryless".

2.3.2 Quantum testers

Quantum testers describe multiple time-step incomplete quantum circuits or networks or test procedures those can be implemented to "measure" or "test" appropriate quantum combs. Prior to describing what general testers are, let us look at the case of tesing a single time-step quantum process and the general test procedures associated with task.

Testing single time-step quantum processes

To learn from or test a single time-step quantum process $\mathcal{E} \in \mathbf{Ch}(\mathcal{H}, \mathcal{K})$ involves test procedures with the following design.

- Prepare a bipartite "test state" $\xi \in \mathbf{St}(\mathcal{H}_{\mathrm{aux}} \otimes \mathcal{H})$, with $\mathcal{H}_{\mathrm{aux}}$ being the auxiliary system.
- Locally implement the to-be-tested channel $\mathcal{E}$ on the probe state, essentially preparing the state $(\mathcal{I} \otimes \mathcal{E})[\xi]$.
- Measure the prepared state with a "test observable" $\mathsf{M} \in \mathbf{Obs}(\underline{n}, \mathcal{H}_{\mathrm{aux}} \otimes \mathcal{K})$. From the observed partial or full statistics, we can learn about the channel.

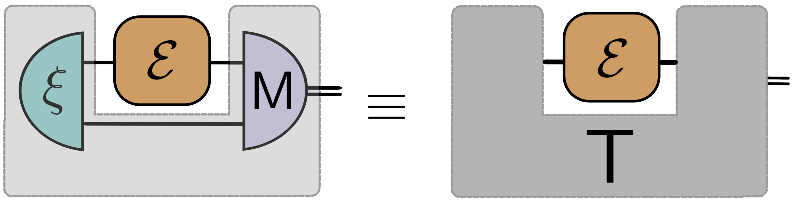

Figure 2.5: The illustration on the left depicts a quantum network, composed of test state ξ and test observable M, that is designed to test the quantum channel $\mathcal{E}$. This testing network described by (ξ, M) can be recast into a single mathematical descriptor T, the quantum tester corresponding to aforementioned pair. This perspective is illustrated on the right.

We can observe here that any such test procedure is prescribed by a pair of test state and test observable $\mathcal{T} = (\xi, \mathsf{M})$. The Born rule describe the statistics produced by the this test procedure as,

$$p_k(\mathcal{T}, \mathcal{E}) = \mathrm{Tr}\left\{\mathsf{M}(k) \cdot (\mathcal{I} \otimes \mathcal{E})[\xi]\right\}. \tag{2.45}$$

It was found by Ziman that this probability distribution can be rewritten as [Zim08],

$$p_k(\mathcal{T}, \mathcal{E}) = \mathrm{Tr}\left\{\mathsf{T}(k)\Phi_{\mathcal{E}}\right\}, \tag{2.46}$$

where $\mathsf{T} \equiv \{\mathsf{T}(k)\}_{k\in\underline{n}}$ is a collection of positive operators on $\mathcal{H}_{\mathrm{aux}} \otimes \mathcal{K}$, those resemble an observable. Referred to as a"process-POVM" by Ziman, this collection of operators, referred to as "process effects", describe the test procedure $\mathcal{T}$, which can be independently viewed as an incomplete circuit with a single open slot (where we "plug in" an appropriate quantum process). Adopting the terminology from [CDP08d] and considering the number of open slots as a characteristic, we refer to such collections as single-slot or 1-slot quantum testers. Similar to quantum observables, these testers also satisfy a normalisation condition which, dissimilar to those observables satisfy, is dependent on the test procedure or specifically the test state. This is given by,

$$\sum_{k\in\underline{n}} \mathsf{T}(k) = \rho \otimes \mathbf{id}_{\mathcal{K}} \quad \text{for some state } \rho \in \mathbf{St}(\mathcal{H}_{\mathrm{aux}}). \tag{2.47}$$

The ρ is referred to as the "normalisation state" of the corresponding tester and is given by $\rho = (\mathrm{Tr}_{\mathcal{H}_{\mathrm{aux}}}\{\xi\})^{\top}$. Moreover, it could happen that two distinct test procedures $\mathcal{T}_1 = (\xi, \mathsf{M})$ and $\mathcal{T}_1 = (\varphi, \mathsf{N})$ could result in the identical statistics for all processes, that is, $p_k(\mathcal{T}_1, \mathcal{E}) = p_k(\mathcal{T}_2, \mathcal{E})$ for all $\mathcal{E}$. Then, Equation 2.46 implies that these two test procedures are described by the same tester T. In this sense, a tester describes equivalent test procedures. Similar to combs, this is also a reason why we are interested in testers; For those scenarios in which we are not curious regarding possible components of any quantum network but the measurements it results in (See Figure 2.5). From now onwards, irrespective of whether we illustrate testing networks with explicit depiction of their components or not, we refer to the testers those networks prescribe and consequently all possible equivalent networks that tester associates itself with.

Testing measurement channels

As an interlude, we will discuss here the particular form of process effects which we can assume without loss of generality, for testers testing measurement channels; This form was presented in [SZ14]. We include this discussion as it shall be particularly used in the investigations later on. As we have seen, the measurement channel associated with an observable $\mathsf{M} \in \mathbf{Obs}(\underline{n}, \mathcal{H})$, with effects $\{M_1, \ldots, M_n\}$ is given by $\Lambda_{\mathsf{M}} \in \mathbf{Ch}(\mathcal{H}, \mathbb{C}^n)$ with corresponding Choi-Jamiołkowski operator $\Phi_{\mathsf{M}} = \sum_k M_k^\top \otimes |k\rangle\langle k|$. Single-slot testers describing the testing of such measurement channels are collections of operators $\{\mathsf{T}_c\} \subset \mathcal{L}(\mathcal{H} \otimes \mathbb{C}^n)$, satisfying $\sum_c \mathsf{T}_c = \xi \otimes \mathbf{id}_{\mathbb{C}^n}$ with ξ being a state on $\mathcal{H}$. We can define projection operators $\pi_k = \mathbf{id}_{\mathcal{H}} \otimes |k\rangle\langle k|$ those which capture the symmetry of the operator Φ_{M} through the identity.

$$\Phi_{\mathsf{M}}^\top = \sum_k \pi_k \Phi_{\mathsf{M}}^\top \pi_k \tag{2.48}$$

Then, during the implementation of the test procedure, the conditional probabilities $p(c|\Lambda_{\mathsf{M}})$ satisfy the identity

$$\begin{aligned}
p(c|\Lambda_{\mathsf{M}}) &= \mathrm{Tr}\left\{\mathsf{T}_c \Phi_{\mathsf{M}}^\top\right\} = \mathrm{Tr}\left\{\mathsf{T}_c \left(\sum_k \pi_k \Phi_{\mathsf{M}}^\top \pi_k\right)\right\} &(2.49)\\
&= \sum_k \mathrm{Tr}\left\{\mathsf{T}_c \pi_k \Phi_{\mathsf{M}}^\top \pi_k\right\} &(2.50)\\
&= \mathrm{Tr}\left\{\pi_k \mathsf{T}_c \pi_k \Phi_{\mathsf{M}}^\top\right\} &(2.51)\\
&= \mathrm{Tr}\left\{\left(\sum_k \pi_k \mathsf{T}_c \pi_k\right) \Phi_{\mathsf{M}}^\top\right\} &(2.52)
\end{aligned}$$

Comparing the first equation and last, we have $\mathsf{T}_c = (\sum_k \pi_k \mathsf{T}_c \pi_k)$. This implies that the test procedure prescribed by the process effects $\{\mathsf{T}_c\}_c$ is experimentally indistinguishable from the test procedure prescribed by the process effects $\{(\sum_k \pi_k \mathsf{T}_c \pi_k)\}_c$. Without loss of generality, we can assume the process effects to be of the form,

$$\mathsf{T}_c = \sum_k H_k^{(c)} \otimes |k\rangle\langle k|\,. \tag{2.53}$$

Looking at the normalisation, we have

$$\sum_c \mathsf{T}_c = \sum_c \sum_k H_k^{(c)} \otimes |k\rangle\langle k| \tag{2.54}$$

$$= \sum_k \left(\sum_c H_k^{(c)} \right) \otimes |k\rangle\langle k| \, . \tag{2.55}$$

Since we are aware of the normalisation of the testers, we require $\sum_k \left(\sum_c H_k^{(c)} \right) \otimes |k\rangle\langle k|$ to be equal to $\xi \otimes \mathbf{id}$. In other words, the summation over k should shift to the second tensor factor as $\sum_k |k\rangle\langle k|$. This is realised when the sum $\left(\sum_c H_k^{(c)} \right)$ does not depend on k. From these, we can read that

$$\sum_c H_k^{(c)} = \xi \quad \text{for all } k. \tag{2.56}$$

Now, the conditional probabilities $p(c|\Lambda)$, after plugging in the expressions for the Choi-Jamiolkowski operator and process effects, we have

$$p(c|\Lambda_{\mathsf{M}}) = \mathrm{Tr} \left\{ \left(\sum_k M_k^{\top} \otimes |k\rangle\langle k| \right) \left(\sum_{k'} H_{k'}^{(c)} \otimes |k'\rangle\langle k'| \right) \right\} \tag{2.57}$$

$$= \delta_{kk'} \mathrm{Tr} \left\{ \sum_k \sum_{k'} M_k^{\top} H_{k'}^{(c)} \otimes |k\rangle\langle k'| \right\} \tag{2.58}$$

$$= \sum_k \mathrm{Tr} \left\{ H_k^{(c)} M_k^{\top} \right\} . \tag{2.59}$$

Quantum testers and their normalisations

After having discussed single-slot testers, we can now generalise the related notion. This notion is generalised when we attempt to describe test procedures designed to test multiple time-step quantum processes. As we have discussed, equivalent classes of such processes are described by quantum combs. So essentially, we want to describe test procedures testing combs with more than one slot. As illustrated in Figure 2.6, a quantum comb with $N-1$ slots is tested when a test procedure $\mathcal{T} = (\tilde{\xi}, \mathcal{C}_1, \ldots, \mathcal{C}_{N-1}, \mathsf{M})$ consisting of appropriate test state $\tilde{\xi}$, test single time-step quantum processes $\mathcal{C}_k$ and test observable M, with outcome set Ω, are used to complete the circuit so as to end up with classical measurement outcomes. Similar to our previous discussions, this test procedure is also collectively described by a single collection of positive operators $\mathsf{T} \equiv \{\mathsf{T}(k)\}_{k\in\Omega}$ and the corresponding incomplete circuit is characterised by N slots. Moreover, such an N-slot tester describes all those equivalent test procedures which are indistinguishable from the measurement statistics. Here also we have a generalised Born rule prescribing the statistics; The statistics produced by an N-slot tester T on an $(N-1)$-slot comb Φ is given by $p_k = \mathrm{Tr}\{\mathsf{T}_k \Phi\}$.

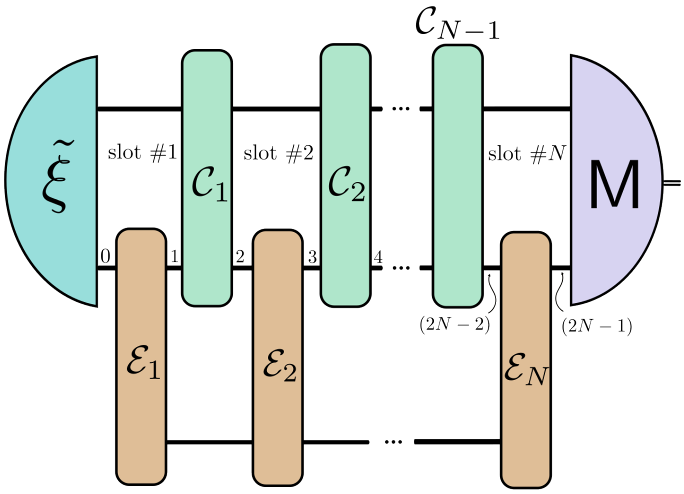

Figure 2.6: An N-slot tester, corresponding to the test procedure $\mathcal{T} = (\tilde{\xi}, \mathcal{C}_1, \ldots, \mathcal{C}_{N-1}, \mathsf{M})$, is testing an $(N-1)$-slot comb.

An N-slot tester T satisfies the following recursive normalisation conditions, which reflect the causal structure of the network as well as the legitimacy of component processes.

$$\sum_k \mathsf{T}(k) = \Theta^{(N)} \otimes \mathbf{id}_{2N-1}, \tag{2.60}$$

$$\mathrm{Tr}_{2N-2}\{\Theta^{(N)}\} = \Theta^{(N-1)} \otimes \mathbf{id}_{2N-3}, \tag{2.61}$$

$$\vdots \tag{2.62}$$

$$\mathrm{Tr}_2\{\Theta^{(2)}\} = \Theta^{(1)} \otimes \mathbf{id}_1 \tag{2.63}$$

$$\mathrm{Tr}\{\Theta^{(1)}\} = 1. \tag{2.64}$$

Here, we refer $\Theta^{(N)}$ as the "normalisation objects" of the corresponding tester. Similar to quantum combs, the operators $\Theta^{(N)}, \Theta^{(N-1)}, \ldots, \Theta^{(1)}$ correspond to subcombs which can be seen as being part of the tester. Given any tester, irrespective of the number of time-steps or open slots it possesses, one can always find a state and observable, on extended spaces, those furnish identical statistics. These observables are of interest to us and are referred to as the "canonical observable" associated to the tester. Given an m-outcome tester with normalisation $\Theta^{(N)} \otimes \mathbf{id}_{2N-1}$, its canonical observable P is constructed through,

$$\mathsf{P}(x) = (\sqrt{\Theta^{(N)}} \otimes \mathbf{id}_{2N-1})\mathsf{T}_x(\sqrt{\Theta^{(N)}} \otimes \mathbf{id}_{2N-1}). \tag{2.65}$$

Chapter 3

Quantum distinguishability

We can infer *quantum distinguishability* as a feature emerging as a result of the quantum information processing task of attempting to distinguish within a finite, known collection of quantum devices of perceived same kind, with each device occuring with a fixed probability. As such, distinguishability can be viewed as a property of any such aforementioned ensemble of devices. The associated information processing task is referred to as *quantum distinguishability task* or *discrimination task*, which is a specific class of quantum hypothesis testing problems. The simplest of these tasks involve distinguishing or discriminating between two given devices, referrred to as binary distinguishability or discrimination tasks.

In this chapter, we start by discussing the structural elements of distinguishability tasks. Rather than how usually distinguishability tasks are discussed, by studying them for states and then generalising to territories of general quantum devices, we adopt a top-down approach. In this approach, we sketch and discuss the tasks in the context of general quantum devices. Then we cast them into the premise of quantum comb distinguishability tasks, which we identify as the most general of these tasks those encompass state, channel, observable distinguishability tasks as special cases. Afterwards, since one major avenue of this thesis revolves around distinguishability tasks of observables, we briefly discuss different distinguishability schemes those can be seen as tailored for distingushing observables; We make use of these schemes in our later investigations into the aforementioned avenue. This chapter concludes with a discussion on anti-distinguishability tasks, which constitute another class of quantum hypothesis testing tasks similar to distinguishability tasks but dissimilar when it comes to the hypotheses involved.

3.1 Architecture of distinguishability tasks

In this section, we discuss the associated aspects of quantum distinguishability tasks.

a priori information and the task

We refer to *a priori* information as the information available or accessible to us, who are going to carry out the distinguishability task. Based on this information, we achieve the capability to design educated discrimination test procedures. As discussed earlier in the introductory passage of this chapter, in any distinguishability task, we are given a quantum device D_x, whose characterisation is unknown to us. Moreover, we are made aware in prior that D_x is one among a well-characterised finite collection of n devices, $D_1, \ldots, D_n$, each of which is known to occur, and thus given to us, with respective probabilities, $p_1, \ldots, p_n$. As such, we can identify the minimal *a priori* information as the ensemble $\mathcal{D}$ made of pairs of devices and their corresponding occurence probabilities, $\mathfrak{D} = \{(D_j, p_j)\}_{j=1}^{n}$. With this information, our task is to guess or identify D_x or equivalently x. Moreover, once we are on an agreement what we are aimed to achieve in a distinguishability task, the ensemble $\mathfrak{D}$ is synonymous with the task, i.e., for the given ensemble.

Test procedures and optimality

Given an ensemble $\mathfrak{D}$, our task is to design a test procedure $\mathcal{T}$ that implements the discrimination task. The task should reflect the distinguishability of $\mathfrak{D}$. Consequently, we expect distinguishability as some function `distinguish` of the ensemble as well as the test procedure, $\texttt{distinguish}[\mathcal{D}, \mathcal{T}]$. Moreover, since we require distinguishability to be a function solely of the ensemble $\mathfrak{D}$, we optimise $\texttt{distinguish}[\mathcal{D}, \mathcal{T}]$ over all test procedures $\mathcal{T}$. The optimised value is the one which is permitted by quantum theory and correspondingly the test procedures achieving this optimality are referred to as optimal test procedures. So, for any distinguishability task, one of the prime goals is to identify optimal test procedures.

Figures of merit

Even though, we wrote `distinguish` as a function reflecting the distinguishabilities of an ensemble when tested with different procedures, the notion itself is ill-defined and was briefly introduced with an intension to eventually convey the notion of optimality. In reality, as we have already discussed, since quantum theory is a probabilistic theory, any quantum distinguishability experiement is characterised by the observed partial or full statistics, which is described by the associated conditional probability distribution. Given an ensemble $\mathfrak{D}$, with n devices, let's say we have a test procedure $\mathcal{T}$ that is designed to have $(n+1)$ decision events or outcomes, $\{1,\ldots,n,?\}$. Here, 1 through n corresponds to respective decisions D_1 through D_n, whereas "?" corresponds to an event where we fail to make any decision. Then the conditional probability distribution given by $p(\mathcal{T}_y|D_x)$ is characteristic of our experiment. We can note that each of these probabilities correspond to three possibilities:

- Success[1] or correct probabilities $p(\mathcal{T}_x|D_x)$ for all x, corresponding to successful events where we identify the device correctly.

- Error probabilities $p(\mathcal{T}_y|D_x)$ with $y \neq x$, corresponding to events where we identify the device erroneously.

- Failure probabilities $p(\mathcal{T}_?|D_x)$ for all x, corresponding to events where we fail to identify or make a decision regarding the tested device.

Now, we can construct figures of merit, which quantify the performance of our distinguishability experiment by appropriately using the above probabilities. This leads us to have the following ones:

- Average success probability $p_s = \sum_x p_x \cdot p(\mathcal{T}_x|D_x)$; As the name suggests, this is quantifies, on average, the success of the distinguishability experiment.

- Average error probability $p_e = \sum_{x,y;x\neq y} p_x \cdot p(\mathcal{T}_y|D_x)$; Similar to the above one, p_e quantifies the experiement's characteristic of producing an erroneous decision on average.

- Total failure probability $p_f = \sum_x p_x \cdot p(\mathcal{T}_?|D_x)$; This quantifies the avaerage probability with which the experiment fails, leading to indecision.

We can note that $p_s + p_e + p_f = 1$. This implies that p_e is a complementary figure of merit to p_s since $p_e = 1 - p_s$, for experiments those do not accomodate any failures.

Strategies

Based on the restrictions we could impose on admittance of errors and/or failures, we can identify the following three strategies, those which are not strictly exclusive from each other.

[1] The qualifier "success" corresponds to the success of the event of correct identification. It does not correspond to the success status of the decision aspect. On the other hand, the failure probabilities corresponds to events of failed decision making.

- Minimum-error tasks. As we have seen, for any given $\mathfrak{D}$, we can write down p_e (or equivalently, p_s). For experiments those seldom admit failures[2], minimum-error distinguishability tasks aim to minimise p_e (or maximise p_s), over all test procedures. Subsequently, these tasks also aim to identify the test procedures those result in this minimisation (or maximisation), which results in p_e^{opt},
$$p_e^{\text{opt}} = \min_{\mathcal{T}} p_e. \tag{3.1}$$

- Perfect unambiguous[3] tasks. These tasks arise when we require the experiment not to accomodate any decision failures as well as any erroneous decisions, i.e., to satisfy the perfect distinguishability conditions $p_f = 0 = p_e$. As such, these tasks essentially ask for the existence of test procedures those can achieve this requirement. We can note that a minimum-error task coincides with a perfect one when p_e^{opt} is found to be zero.

- Imperfect unambiguous tasks. These tasks arise when an ensemble does not admit iteself to be perfectly unambiguously distinguished and when we require or design the experiment not to accomodate any erroneous decisions, but at a potential cost of having decision failures, i.e., to satisfy the "imperfect unambiguous" no-error conditions $p_e = 0$ and $p_f \neq 0$. Thus, we can be sure that if the experiment run did not fail, the decision made is correct for sure. Similar to perfect tasks, these tasks ask for the existence of test procedures those can achieve this requirement. Once such a collection is identified, then the subsequent task is to find a procedure that minimises p_f over this collection of test procedures.

3.2 Distinguishing quantum combs

We assert that all quantum distinguishability tasks, within the scope of this thesis[4], can be considered a quantum comb distinguishability task, essentially claiming that $\mathfrak{D}$ is always an ensemble of combs. This identification is made as a consequence of the observation that all quantum devices can be viewed as a quantum comb with appropriate number of open slots. This enables us to recast our previous discussion for quantum combs. Let $\mathfrak{D} = \{(\Phi_k^{(N)}, p_k)\}_{k=1}^m$ be an ensemble of m quantum combs with N slots. The distinguishability experiments of this ensemble is described by quantum testers $\mathsf{T}^{(N+1)}$ with $(N+1)$ slots. These

[2]Even though it is possible to study minimum-error tasks with some fixed non-zero p_f, in this thesis it is not considered.

[3]We shall omit the qualifier "unambiguous" and just refer to these tasks as perfect ones.

[4]restricted to quantum processes with definite causal internal structures.

testers have m process effects T_j each corresponding to the decision event of identifying the comb $\Phi_j^{(N)}$. Moreover, for imperfect unambiguous experiments, there will be an extra process effect $\mathsf{T}_?$ corresponding to the event of decision failure. Regardless, each of these experiments are characterised by the conditional probability distribution $p(\Phi_y^{(N)}|\Phi_x^{(N)})$, which is given by

$$p(\Phi_y^{(N)}|\Phi_x^{(N)}) = \mathrm{Tr}\left\{\Phi_x^{(N)}\mathsf{T}_y\right\}. \tag{3.2}$$

From this distribution, the aforementioned figures of merit can be constructed and subsequent distinguishability studies can be carried out.

3.2.1 Perfect distinguishability

Two combs, regardless of the probabilities with which they occur within an ensemble, can be perfectly distingusihed if they satisfy the condition contained in the following proposition.

Proposition 3.2.1. Two N-slot quantum combs Φ_1, Φ_1 can be perfectly discriminated, there exists at least one appropriate binary $(N+1)$-slot quantum tester T with normalisation $\Xi \otimes \mathbf{id}$ such that

$$\Phi_1(\Xi \otimes \mathbf{id})\Phi_2 = O. \tag{3.3}$$

Proof. Perfect discrimination of two N-slot combs Φ_1 and Φ_2 is admitted if there exists a binary tester T_1, T_2 with normalisation $\Xi \otimes \mathbf{id}$, satisfying the conditions $\mathrm{Tr}[\Phi_1\mathsf{T}_2] = 0$ and $\mathrm{Tr}[\Phi_2\mathsf{T}_1] = 0$, those required to be simultaneously satisfied. This translates to the operator equalities $\Phi_1\mathsf{T}_2 = O$ and, due to cyclicity of trace, $\mathsf{T}_1\Phi_2 = O$. Right multiplying the first equation with Φ_2 and left multiplying the second equation with R_1, we have $\Phi_1\mathsf{T}_2\Phi_2 = O$ and $\Phi_1\mathsf{T}_1\Phi_2 = O$, respectively. Adding these two equalities, we have $\Phi_2(\mathsf{T}_1 + \mathsf{T}_2)\Phi_1 = O$, which leads to $\Phi_1(\Xi \otimes \mathbf{id})\Phi_2 = O$. □

Corollary 3.2.1. An ensemble of two quantum states $\{(\rho_1, \frac{1}{2}), (\rho_2, \frac{1}{2})\}$ is perfectly distinguishable if and only if

$$\rho_1\rho_2 = O. \tag{3.4}$$

This implies that the two states should have orthogonal supports.

Corollary 3.2.2. An ensemble of two quantum channels $\{(\mathcal{E}_1, \frac{1}{2}), (\mathcal{E}_2, \frac{1}{2})\}$ is perfectly distinguishable if there exists a binary 1-slot tester with normalisation $\xi\otimes\mathbf{id}$ satisfying the condition

$$\Phi_{\mathcal{E}_1}(\xi \otimes \mathbf{id})\Phi_{\mathcal{E}_2} = O. \tag{3.5}$$

Perfectly distinguishing more combs

For an arbitrary ensemble of combs $\{(\Phi_j, p_j)\}_{j=1}^n$, with respect to the number of combs belonging to it, there could be more than two combs, that is, $n > 2$. In such cases, each comb should be perfectly distinguishable from the rest. Consequently, perfect distinguishability of such an ensemble translates to satisfying the following $\binom{n}{2}$ operator equalities for all $i, j \in \underline{n}$ with $i \neq j$,

$$\Phi_{\mathcal{E}_i}(\xi \otimes \mathbf{id})\Phi_{\mathcal{E}_j} = O. \tag{3.6}$$

3.2.2 Minimum-error distinguishability

As we have discussed, minimum-error distinguishability task for the given ensemble of combs aims at finding appropriate testers those minimise the average error probability to get p_e^{opt},

$$p_e^{\text{opt}} = \min_{\mathsf{T}} \sum_{x,y;x\neq y} p_x \mathrm{Tr}\left\{\Phi_x^{(N)} \mathsf{T}_y\right\}. \tag{3.7}$$

Operational distances

The above optimisation, when carried out for the binary distinguishability task of a specific quantum device class when the devices appear with equal probabilities, results in identifying an emergent operational[5] distance for the device class. Let us look some well established ones here.

- For the case of states, given an ensemble $\{(\rho_1, \frac{1}{2}), (\rho_2, \frac{1}{2})\}$, the optimal error probability is found to be [Hel69]
$$p_e^{\text{opt}} = \frac{1}{2}\left(1 - \|\rho_1 - \rho_2\|_1\right). \tag{3.8}$$
Here, we can identify that the operator trace norm (given by the sum of absolute value of eigenvalues of the operator) quantity $\|\rho_1 - \rho_2\|_1$ as the operational distance between the two states ρ_1 and ρ_2.

- For the case of channels, given an ensemble $\{(\mathcal{E}_1, p_1), (\mathcal{E}_2, p_2)\}$, the optimal error probability is given by [Sac05, Rag02]
$$p_e^{\text{opt}} = \frac{1}{2}(1 - \frac{1}{2}\|\mathcal{E}_1 - \mathcal{E}_2\|_\diamond). \tag{3.9}$$
Here, we can identify that operational distance between the two aforementioned channels is given by the completely bounded norm [Pau02] or diamond norm $\|\mathcal{E}_1 - \mathcal{E}_2\|_\diamond$ which is defined as $\|\mathcal{B}\|_\diamond := \max_{d,\xi} \|(\mathcal{I}_d \otimes \mathcal{B})(\xi)\|_1$.

[5] with respect to the operational task of distinguishability.

– For the general case of quantum combs, given an ensemble N-slot combs $\{(\Phi_1, \frac{1}{2}), (\Phi_2, \frac{1}{2})\}$, the optimal error probability is given by [CDP09]

$$p_e^{\mathrm{opt}} = \frac{1}{2}\left(1 - \frac{1}{2}\max_T \left\| \sqrt{T^\top}(\Phi_1 - \Phi_2)\sqrt{T^\top} \right\|_1 \right). \tag{3.10}$$

Here, we can identify $\frac{1}{2}\max_T \left\| \sqrt{T^\top}(\Phi_1 - \Phi_2)\sqrt{T^\top} \right\|_1$ as the operational distance between the two combs, where the maximisation is done over all possible normalisation of $(N+1)$-slot testers T.

3.2.3 Imperfect unambiguous distinguishability

As discussed earlier, when we find that a particular ensemble of combs $\{(\Phi_j, p_j)\}_{j=1}^n$ is not perfectly distinguishable, that is there does not exist an approriate tester that can achieve perfect distinguishability, we can shift our focus to imperfectly and unambiguously distinguish the ensemble. To achieve this task, there should exist a tester T with $N+1$ outcomes, with one outcome corresponding to failure and described by the non-zero process effect $\mathsf{T}_?$, that satisfies the following "imperfect unambiguous" no-error conditions,

$$\mathrm{Tr}\left\{\Phi_x \mathsf{T}_y\right\} = 0; \quad \forall x \neq y \quad \text{with} \quad x, y \in \underline{n}. \tag{3.11}$$

Then, the average failure probability of the experiment is given by

$$p_f = \sum_{x=1}^{n} p_x \mathrm{Tr}\left\{\Phi_x \mathsf{T}_?\right\}. \tag{3.12}$$

Our task is then to minimise this failure probability by optimising over quantum testers satisfying the conditions 3.11. Moreover, if there does not exist any tester satisfying these conditions to begin with, then we assert that the given ensemble cannot be distinguished imperfect-unambiguously.

3.3 Distinguishing quantum observables

As briefly discussed in the introductory chapter, since the task of labeling an observable which we introduce and investigate in this thesis is related to the more general task of observable distinguishability, we restrict the following discussion on distinguishability networks or schemes to those which are used in distinguishing observables; These schemes are those testing procedures or networks which test a given observable so as to distinguish it from the rest of the observables from the ensemble.

Number of shots of the observable

Given an ensemble of observables $\{(\mathsf{M}_j, p_j)\}_{j=1}^n$, we can construct schemes based on how many uses or implementations or queries of the chosen observable is available, either by virtue of the availability of the measurement device implementing the observable or self-restriction, for the investigation. If each of the observable from the ensemble can occur with N queries, then essentially this means that the ensemble to be distinguished is $\{(\mathsf{M}_j^{\otimes N}, p_j)\}_{j=1}^n$. We refer to this number as the number of shots associated with the distinguishability task. This enables us to classify the distinguishability schemes as belonging to two regimes: single-shot (when $N = 1$) or multiple-shot (when $N > 1$).

3.3.1 Single-shot schemes

In this subsection, we classify and discuss distinguishability schemes when a single use or query of the chosen observable is available. Depending on the scheme and how we want to analyse the task, the ensemble can be considered equivlanelty as either $\{(\mathsf{M}_j, p_j)\}_{j=1}^n$ or $\{(\Phi_j, p_j)\}_{j=1}^n$, where Φ_j is the Choi-Jamiołkowski operator of the measurment channel Λ_{M_j}.

Simple schemes

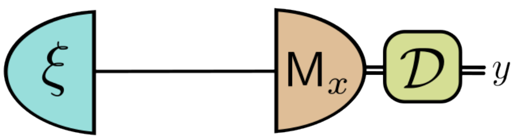

Figure 3.1: Single-shot simple scheme distinguishing an ensemble $\{(\mathsf{M}_j, p_j)\}_{j=1}^n$; M_x is the chosen unknown observable and one among the possible ones. Based on the ensemble, a state ξ is prepared and the observable is measured on this state. $\mathcal{D}$ is the final processing we implement conditioned on the classical information we receive from the measurement process. This post-processing eventually leads to a decision y or M_y regarding the chosen observable.

As depicted in Figure 3.1, while in the single-shot scenario, through a simple scheme we measure the chosen observable on an appropriately prepared probe state. This measurement process furnishes some classical information, conditioned on which some post-processing is carried out to arrive at a decision regarding the chosen observable. While implementing such schemes, our goal is to find the optimal probe state so as to maximise the distinguishability.

Assisted schemes

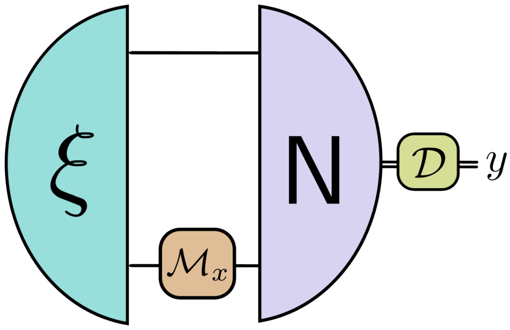

Figure 3.2: In a single-shot assisted scheme, the measurement channel corresponding to the chosen observable $\mathcal{M}_x$ is implemented locally on an entangled bipartite probe state ξ. The transformed bipartite state is then measured by a final probe observable N. From the obtained classical information, we can make decisions regarding the chosen observable. Here, we can note that the pair (ξ, N) corresponds to a 1-slot tester.

Figure 3.2 depicts the scenario of entanglement-assisted single shot distinguishability schemes, where entanglement is used to assist the task so as to potentially improve its performance. Without loss of generality, the unknown observable M_x is modeled as its corresponding measurement channel $\mathcal{M}_x = \Lambda_{\mathsf{M}_x}$. This channel is implemented locally on a, without loss of generality, bipartite system which is prepared in an entangled probe state ξ. Then, the transformed state $\mathcal{I} \otimes \mathcal{E}(\xi)$ is then measured by another probe observable N whose classical outcome is post-processed to arrive at a decision regarding the choice. We can note that such an assisted testing procedure, comprised of the test state objects (ξ, N), is equivalently described by a binary 1-slot quantum tester T. In such cases, the post-processing is part of the tester's measurement events. Moreover, we can also note that the simple scheme described above can also be described by 1-slot quantum testers.

3.3.2 Multiple-shot schemes

In this section, we classify distinguishability schemes when more than one use or query of the chosen observable is available. This is when we distinguish, in effect, the ensemble $\{(\mathsf{M}_j^{\otimes N}, p_j)\}_{j=1}^n$ with $N > 1$. The corresponding quantum comb to these N copies of the observable is given by "link multiplying" the component measurement channels Φ_{M}. Since there are no quantum memories connecting these copies, this $(N-1)$-slot quantum comb is given by tensor product of Choi-Jamiołkowski operators of the channels,

$$\underbrace{\Phi_{\mathsf{M}} * \cdots * \Phi_{\mathsf{M}}}_{N \text{copies}} = \Phi_{\mathsf{M}} \otimes \cdots \otimes \Phi_{\mathsf{M}} = \Phi_{\mathsf{M}}^{\otimes N}. \tag{3.13}$$

As such, when we investigate this problem as a comb discrimination problem, we are essentially distinguishing the ensemble $\{(\Phi_{\mathsf{M}_j}^{\otimes N}, p_j)\}_{j=1}^{n}$. Moreover, as we shall see, these schemes are richer than single-shot ones in terms of the additional possibilities accommodated by them. In these schemes, improvement in distinguishability not only comes from the multiple copies or queries available but also from the "adaptivity" of the networks as discussed below.

Simple schemes

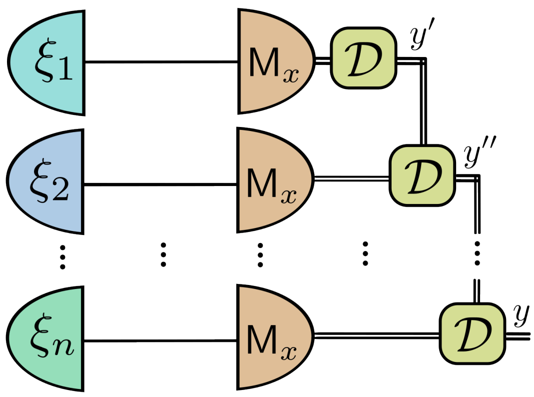

Figure 3.3: We can see that this multiple-shot simple scheme is in fact a collection of single-shot simple schemes, each with a different probe state ξ_j. If all of these states are same, then this scheme is equivalent to its single-shot version as one would expect.

As illustrated here, multiple-shot simple schemes consists of several single-shot simple schemes, those which necessarily do not have any particular ordering in time and are independent of each other. We can also note that, similar to the single-shot scenario, in the multi-shot scenario there is no entanglement-assistance. Eventhough these schemes are in general not relevant for observable distinguishability tasks, they hold importance for our investigations into the labeling tasks.

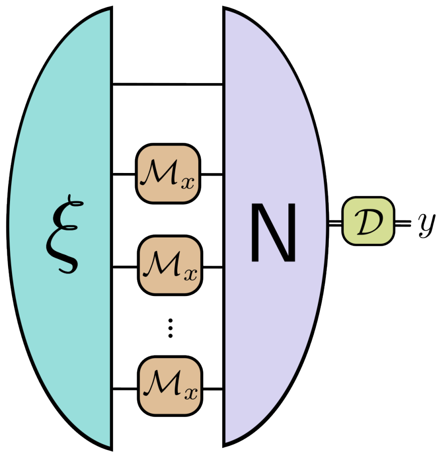

Figure 3.4: In the parallel scheme illustrated here when you have access to N queries of the chosen observables, an $N+1$-partite quantum state is prepared in the state ξ. Subsequently, the measurement channels are implemented locally as shown here. The transformed state $(\mathcal{I} \otimes \mathcal{M}^{\otimes N})[\xi]$ is then measured by the global observable N, resulting in a decision. The pair (ξ, N) corresponds to a 1-slot quantum tester here as well.

Parallel schemes

Parallel schemes are generalised versions of the single-shot assisted schemes. In such schemes, when N queries of the measurement channel are available, each channel is implemented locally on an appropriate $N+1$-partite system prepared in an entangled probe state as shown in Figure 3.4. The resultant state after these transformations are then measured using a global observable to arrive at decisions regarding the chosen observable. Moreover, such schemes are also described by binary 1-slot quantum testers with N outcomes.

Adaptive schemes

Adaptive schemes, one such illustrated in Figure 3.5, harness the most out of the availability of multiple queries of the chosen observable, with respect to improvement in performances of the distinguishability task. This is because, on top of being assisted by entanglement, intermediate processing transformations are implemented in between the implementation of the measurement channels across in time. Moreover, this means that part of the probe state goes through one query of the measurement channel, followed by the processing channel and then to another query of the memory channel and so on. As such, the memory effects ramified from this procedure could potentially result in improvement in distinguishability, when compared to the previous schemes. This feature of memory is a characteristic of adaptive schemes. Now, when N copies or queries of the chosen observable is available, that is, while being in the N-shot scenario, the corresponding adaptive schemes are described by N-slot quantum testers.

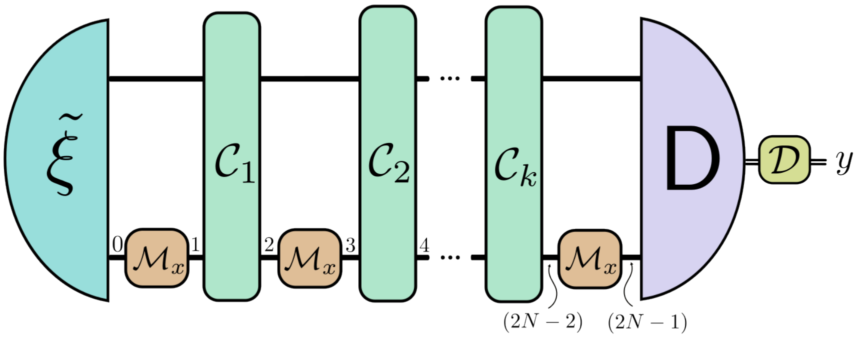

Figure 3.5: An adaptive scheme distinguishing the chosen observable when its corresponding measurement channel $\mathcal{M}$ is available for N queries. The scheme, comprised for the probe state $\tilde{\xi}$, intermediary probe transformations $\mathcal{C}_j$ and final probe observable D, is described by an appropriate N-slot quantum tester.

Chapter 4

Single-shot labeling of quantum observables

In this chapter, we discuss the notion of an unlabeled quantum observable to motivate and introduce the task of labeling of such an observable. Subsequently, this task is identified as a specific distinguishability problem of observables wherein which the associated ensemble possesses a specific symmetry. After the introduction of the task, rest of the chapter deals with investigations on single-shot labeling, where we attempt to perform this task when a single copy or implementation of associated observables are accessible.

Revisiting the observable

Prior to identifying the task of labeling, let us revisit the notion of an observable and pay more attention to its aspects. We have seen that an observable $\mathsf{M} \in \mathbf{Obs}(\Omega, \mathcal{H})$, associated to a quantum measurement, is a specific prescription for assigning each and every outcome label $\omega \in \Omega$ to some effects $E \in \mathbf{Eff}(\mathcal{H})$. As such, we can assert that the "complete" description of an observable is given by the following three sets of objects:

- the set of outcome labels $\{\omega\}$
- the set of effects $\{E\} \subset \mathbf{Eff}(\mathcal{H})$ or a specific resolution of the $\mathbf{id}_{\mathcal{H}}$
- the set of prescriptions $\{\omega \mapsto E\}$, which dictate the association between outcome labels and their corresponding effects. We refer each such prescription as a "outcome-effect association" or "labeling".

As a circular observation, we can see that the associated normalised-effect-valued-measure M is an embodiment of these descriptors.

Unlabeled observables

Now, we can have observables with incomplete description, with the information regarding the prescriptions $\{\omega \mapsto E\}$ being missing or lost. We refer to such observables as "unlabeled" observables; In this context, "labeled" observables are those with complete descriptions. As such, an unlabeled observable has only two descriptors, the outcome labels $\{\omega\}$ and an $\mathbf{id}_{\mathcal{H}}$ resolution as the effects $\{E\}$. In fact, an unlabeled observable is indistinguishable, as such, from an labeled observable whose range is equal to the collection of effects associated with the unlabeled one. We can now formalise what we mean by an unlabeled observable by noticing the following equivalence relation. Two labeled observables $\mathsf{A}, \mathsf{B} \in \mathbf{Obs}(\underline{n}, \mathcal{H})$ are equivalent to each other, $\mathsf{A} \equiv \mathsf{B}$, if there exists a permutation $\pi : \underline{n} \to \underline{n}$ such that $\mathsf{A}(j) = \mathsf{B}(\pi(j))$ for all $j \in \underline{n}$. As such, given the quantum system $\mathcal{H}$, its set of observables with n outcomes[1], $\mathbf{Obs}(\underline{n}, \mathcal{H})$, is partitioned into equivalent classes,

$$[\mathsf{M}] := \{\mathsf{M}' \in \mathbf{Obs}(\underline{n}, \mathcal{H}) : \mathsf{M} \equiv \mathsf{M}'\}. \tag{4.1}$$

With this notion, when we have an unlabeled observable we can equivalently treat it as one among the elements of the appropriate equivalent class. The cardinality of a particular class $[\mathsf{M}]$, or equivalently the number of non-identical labeled observables included in $[\mathsf{M}]$, depends upon the multiplicity of the effects involved. If we look at the case where all the n effects are identical to each other[2], then there is no non-identical permutation and consequently $[\mathsf{M}]$ is singleton. On the other hand, when all the involved effects are non-identical to each other, there are $n!$ permutations leading to $n!$ non-identical observables included in $[\mathsf{M}]$. When there are some effects those appear with multiplicities more than 1, the number of non-identical observables are between 1 and $n!$, depending upon the number of effects appearing more than once.

Let us have an illustration of these notions. Given a labeled observable $\mathsf{M}_1 \in \mathbf{Obs}(\underline{3}, \mathcal{H})$ with non-identical effects, we can construct the corresponding equivalence class $[\mathsf{M}_1] = \{\mathsf{M}_1, \mathsf{M}_2, \mathsf{M}_3, \mathsf{M}_4, \mathsf{M}_5, \mathsf{M}_6\}$. If the composing effects are E_1, E_2, E_3, then these observables are described by the following outcome-effect associations.

[1] We should note that we assume each of the n outcomes to have a non-zero probability of occuring in some experiment. Essentially, they should not be associated with zero effects O; Otherwise we can consider and add as many outcomes to the description of an observable and have any value for $\underline{n}$.

[2] This essentially implies that each of the effects is $\frac{1}{n}\mathbf{id}_{\mathcal{H}}$.

$$\begin{aligned}
&\mathsf{M}_1(1) = E_1 \qquad && \mathsf{M}_2(1) = E_1 \qquad && \mathsf{M}_3(1) = E_2 \\
&\mathsf{M}_1(2) = E_2 && \mathsf{M}_2(2) = E_3 && \mathsf{M}_3(2) = E_1 \\
&\mathsf{M}_1(3) = E_3 && \mathsf{M}_2(3) = E_2 && \mathsf{M}_3(3) = E_3
\end{aligned} \tag{4.2}$$

$$\begin{aligned}
&\mathsf{M}_4(1) = E_2 \qquad && \mathsf{M}_5(1) = E_3 \qquad && \mathsf{M}_6(1) = E_3 \\
&\mathsf{M}_4(2) = E_3 && \mathsf{M}_5(2) = E_1 && \mathsf{M}_6(2) = E_2 \\
&\mathsf{M}_4(3) = E_1 && \mathsf{M}_5(3) = E_2 && \mathsf{M}_6(3) = E_1
\end{aligned} \tag{4.3}$$

Now, when we are given an unlabeled observable M_x composed of the same effect operators E_1, E_2 and E_3, we can conclude that it is one among the above six observables, that is, $\mathsf{M}_x \in [\mathsf{M}_1]$.

We can also clarify furthur the notion by having the following discussion with ideal Stern-Gerlach measurement devices. To recall, such devices implement sharp binary observables those belong to $\mathbf{Obs}(\underline{2}, \mathbb{C}^2)$. If the spatial direction associated with a Stern-Gerlach device is oriented along the $+z$, the observable it implements $\mathsf{S}_{\uparrow;z}$ is given by the effects,

$$\mathsf{S}_{\uparrow;z}(1) = |{\uparrow_z}\rangle\langle{\uparrow_z}| \quad \text{and} \quad \mathsf{S}_{\uparrow;z}(2) = |{\downarrow_z}\rangle\langle{\downarrow_z}|\,. \tag{4.4}$$

Here, $|{\uparrow_z}\rangle\langle{\uparrow_z}|$ corresponds to the event when a measurement recording is observed in the $+z$ direction and $|{\downarrow_z}\rangle\langle{\downarrow_z}|$ is complementarily described. Now, when the orientation is along the $-z$ axis, the associated observable $\mathsf{S}_{\downarrow;z}$ is given by the effects,

$$\mathsf{S}_{\downarrow;z}(1) = |{\downarrow_z}\rangle\langle{\downarrow_z}| \quad \text{and} \quad \mathsf{S}_{\downarrow;z}(2) = |{\uparrow_z}\rangle\langle{\uparrow_z}|\,. \tag{4.5}$$

We can see that even though the above two observables describe strictly two different measurement scenarios, they are composed of the same (non-ordered) collection of effects and by our previous discussion they are equivalent to each other.

From one labeling to another: Relabel post-processing

To recall, given an observable $\mathsf{M} \in \mathbf{Obs}(\Omega, \mathcal{H})$, composed of effects $M_1, \dots M_n$, the associated measurement channel Λ_{M} is given by its action as

$$\Lambda_{\mathsf{M}}(\varrho) = \sum_k \mathrm{Tr}\{\varrho M_k\} |x_k\rangle\langle x_k|\,, \tag{4.6}$$

where $\{|x_k\rangle\}$ forms an orthonormal basis for the system $\mathbb{C}^n$. As discussed, given such an observable it is straightforward to write down rest of the observables included in the class $[\mathsf{M}]$ by permuting the effect assignments to the labels. Now, when a measurement channel is given we can construct the measurement channels of rest of the observables by introducing permutation operators $P_\sigma := \sum_k |x_{\sigma(k)}\rangle\langle x_k|$, were σ are permutations on the label set Ω. Then another measurement channel $\Lambda_{\mathsf{M}_\sigma}$ is arrived at through $\Lambda_{\mathsf{M}_\sigma}(\varrho) = P_\sigma \Lambda_{\mathsf{M}} P_\sigma^\dagger$. These conversions are correspondingly reflected in their Choi-Jamiołkowski operators as $\Phi_{\mathsf{M}_\sigma} = (\mathbf{id} \otimes P_\sigma)\Phi_{\mathsf{M}}(\mathbf{id} \otimes P_\sigma)^\dagger$.

In the previous passage, we have discussed equivalent ways through which we can construct the whole equivalence class of observables, associated with an unlabeled observable. We should note that these constructions do not correspond directly to any physical processes. On the other hand, we can identify that the physical processes, although classical in nature, associated with converting an observable to another "relabeled" one, are specific observable postprocessings referred to as "relabelings" [HHP12]. During the relabeling process, if we are not associating a single label to different outcomes, given a quantum observable $\mathsf{M} \in \mathbf{Obs}(\underline{n}, \mathcal{H})$ with not all effects being identifical to each other, we can have anywhere between two to $n!$ (depending on other effect multiplicities) relabled observables through post-processings $\{\delta^k\}$ with $\mathsf{M}_k(x) = \sum_y \delta^k_{xy}\mathsf{M}(y)$. We can note that the entries to these postprocessing matrices are either 0 or 1. See Figure 4.1 for an illustration of relabeling.

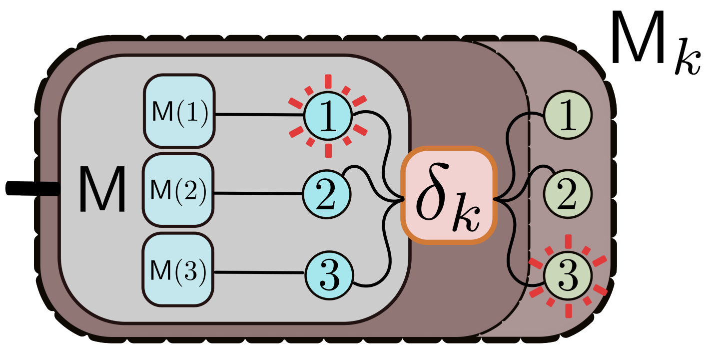

Figure 4.1: An illustration of a measurement device that implementes a relabeling of the observable M as M_k, due to the postprocessing mediated by δ_k. We can see how the label "1" of the M is relabeled (through internal reconnections) to "3" of M_k; When label 1 is recorded for M, correspondingly 3 is recorded for M_k.

4.1 The labeling task

To label an unlabeled observable

As we have discussed earlier, an unlabeled observable is one whose label-effect associations are lost. As such, given such an observable we can ask the question whether we can find this lost information, in full or partially, by performing some experiments. In this thesis, such experiments are said to "label" the unlabeled observables.

Distinguishing different labelings

A ramified task related to labeling unlabeled observables is to identify an unlabeled observable, or equivalently its labeling[3], when the observable has been made available by random sampling from the total collection of labeled observables $[\mathsf{M}]$. In fact, labeling an unlabeled observable is equivalent to this specific observble distinguishability task where we assume each of labeled observables from $[\mathsf{M}]$ occur with equal probability. This is because when we distinguish a particular observable from the rest we are essentially identifying the labeling of the observable.

We can now formally sketch this distinguishability task. Given an unlabeled observable, hailing from the class $[\mathsf{M}]$, the labeling problem or task is the disinguishability problem for the collection of observables $[\mathsf{M}]$ where the prior probabilities with which each of labeled observables occur is $\frac{1}{k}$, where k is the cardinality of the set $[\mathsf{M}]$. Since these tasks are observable distinguishability tasks, they can be recast into channel distinguishability tasks, where the corresponding measure-and-prepare channels of the involved observables are distinguished.

In the rest of this thesis, we will omit from using the qualifier "unlabeled" in front of observables when it is clear from the context that they are unlabeled and being studied for labeling tasks. Moreover, within such labeling contexts, when we write statements having "an observable $\mathsf{M} \in \mathbf{Obs}(\underline{n}, \mathcal{H})$" we necessarily refer to an unlabeled observable where M being the representative of the class $[\mathsf{M}]$. The cases where $[\mathsf{M}]$ is singleton, those cases where all the involved effects are identical to each other, there does not exist a distinguishability problem. As such, in the rest of the thesis the results discovered does not take into account these pathological cases.

[3] In this context *labeling* is used as a noun.

Single-shot labeling tasks

In this chapter, we study the labeling task in the single-shot regime, that is, when we access to only a single copy or use of the measurement device implementing the unlabeled observable.

4.2 Binary observables

We begin the study of single-shot labeling by exploring the case of binary observables on arbitrary quantum systems $\mathcal{H}$, $\mathsf{M} \in \mathbf{Obs}(\underline{2}, \mathcal{H})$. Labeling of a binary observable is similar to the basic and archetypical case of distinguishing two states. This is because, for the non-trivial case with non-identical effects, there are two non-identical labeled observables. Then, in the corresponding labeling problem we require to distinguish between the following observables, M and N, composed of effects $M_1, M_2 \in \mathbf{Eff}(\mathcal{H})$:

$$\begin{aligned} \mathsf{M}(1) &= M_1 \qquad & \mathsf{N}(1) &= M_2 \\ \mathsf{M}(2) &= M_2 \qquad & \mathsf{N}(2) &= M_1 \end{aligned}$$

The measurement channels, $\Lambda_{\mathsf{M}}, \Lambda_{\mathsf{N}} \in \mathbf{Ch}(\mathcal{H}, \mathbb{C}^2)$, of these observables are given by their actions as

$$\Lambda_{\mathsf{M}}(\rho) = \operatorname{Tr}\{M_1\rho\}|0\rangle\langle 0| + \operatorname{Tr}\{M_2\rho\}|1\rangle\langle 1| \tag{4.7}$$
$$\Lambda_{\mathsf{N}}(\rho) = \operatorname{Tr}\{M_2\rho\}|0\rangle\langle 0| + \operatorname{Tr}\{M_1\rho\}|1\rangle\langle 1|\,. \tag{4.8}$$

Correspondingly, let Φ_{M} and Φ_{N} be the Choi operators of $\Lambda_{\mathsf{M}}, \Lambda_{\mathsf{N}}$, respectively. These are given by

$$\Phi_{\mathsf{M}} = M_1^{\top} \otimes |0\rangle\langle 0| + M_2^{\top} \otimes |1\rangle\langle 1| \tag{4.9}$$
$$\Phi_{\mathsf{N}} = M_2^{\top} \otimes |0\rangle\langle 0| + M_1^{\top} \otimes |1\rangle\langle 1|\,. \tag{4.10}$$

4.2.1 Perfect labeling

In perfect labeling, we are interested to know which observables admit to be perfectly labeled in a single shot. Our subsequent investigation results in the following theorem.

Theorem 4.2.1. A binary observable $\mathsf{M} \in \mathbf{Obs}(\underline{2}, \mathcal{H})$ can be *perfectly labeled* in a single shot if and only if atleast one of the composing non-identical effects is *rank deficient*.

Proof. Proposition 3.2.1 furnishes the perfect distinguishability condition for two quantum combs. For our specific case of discriminating two measurement channels, Λ_{M} and Λ_{N}, the proposition translates to the existence of a binary 1-slot tester T, with normalisation $\xi \otimes \mathbf{id}$, satisfying the condition

$$\Phi_{\mathsf{M}}(\xi \otimes \mathbf{id})\Phi_{\mathsf{N}} = O. \tag{4.11}$$

Plugging in the expressions for the Choi operators into the above equations results in,

$$M_1^{\top} \xi M_2^{\top} \otimes |0\rangle\langle 0| + M_2^{\top} \xi M_1^{\top} \otimes |1\rangle\langle 1| = O. \tag{4.12}$$

From this, we can read $M_1^{\top} \xi M_2^{\top} = O$ and $M_2^{\top} \xi M_1^{\top} = O$. Let us take the following from this.

$$M_1 \xi M_2 = O. \tag{4.13}$$

Now, suppose M is a full rank observable, such that inverses of the effects, M_1^{-1} and M_2^{-1}, exist. We can then sandwich the above equation with these inverses as,

$$M_1^{-1}(M_1 \xi M_2)M_2^{-1} = O \tag{4.14}$$

$$\xi = O. \tag{4.15}$$

Since no quantum state ξ can be the zero operator, by contradiction this reveals that there does not exist any binary 1-slot tester T that can perfectly label a full-rank binary obseravble. On the contrary, suppose that one of the underlying effects, say M_1 is rank deficient, that is, there exists at least one state $|\psi\rangle \in \mathcal{H}$ such that $M_1 |\psi\rangle = 0$. Then, the perfect distinguishability condition is satisfied by 1-slot testers having normalisation state $\xi = |\psi\rangle\langle\psi|$, since $M_1 |\psi\rangle\langle\psi| M_2 = O$. □

As a no-go theorem

We can equivalently rephrase Theorem 4.2.1 as a no-go statement: "If a binary observable is a full-rank observable, then it does not admit perfect labeling in a single shot".

Entanglement assistance is unnecessary

As discussed in chapter 2, 1-slot testers describe the most general test procedures for testing quantum channels characterised by a single time step. Alongside, we found that perfect labeling is possible if and only if one of the underlying effects of the binary obseravble is rank deficient. That is, one of the effects should have eigenvalue 0, at least with a multiplicity of one. Moreover, any tester with normalisation state corresponding to a pure state hailing from those corresponding null spaces can effectively result in perfect labeling and pure normalisation states correspond to test procedures without any entanglement assistance.

Simple scheme is suffcient

Suppose M_1, one of the two effects is rank deficient. Correspondingly M_2 has an eigenvalue 1 with at least a multiplicity of one. That is, there exists a pure state $|\varphi\rangle$ such that $M_2 |\varphi\rangle = 0$. As such, we can device a simple scheme as follows. Prepare the state $|\varphi\rangle\langle\varphi|$ and measure the unlabeled observable on this state. The outcome label that gets registered corresponds to the effect M_2. This also means that the other label corresponds to the effect M_1. This essentially shows that such a simple scheme, without entanglement-assistance, is suffcient for perfect labeling in a single shot.

Proposition 4.2.1. If a binary observable M does not admit perfect labeling, then there does not exist any pre-processing quantum channel that can lead to perfect labeling.

Let $\mathsf{M} \in \mathbf{Obs}(\underline{2}, \mathcal{H})$ be a binary observable that does not admit perfect labeling and consequently composed of full-rank observables. In the simple scheme this translates to fact that there does not exist any probe state with which we can perfectly label this unlabeled observable in a single-shot. On the other hand, let us assume that without loss of generality $\mathcal{E} \in \mathbf{Ch}(\mathcal{H}, \mathcal{H})$ is a pre-processing quantum channel that could transform the observable to a rank-deficient one, that is, $\mathcal{E}^*(\mathsf{M})$ is rank-deficient. This would imply that there exists at least one state $|\varphi\rangle$ with which we can label the transformed observable. This scenario is equivalent to testing the observable M with the state $\mathcal{E}(|\varphi\rangle\langle\varphi|)$ and resulting in perfect labeling, which contradicts our initial assumption regarding M.

4.2.2 Minimum-error labeling

In single-shot minimum-error labeling, given the unlabeled observable M, our goal is to design labeling experiments which optimise the performance of the task through minimisation of the average error probability associated with labeling it or distinguishing between M and N. In order to achieve this we shall first minimise this quantity which will subsequently enable us to identify optimal testers and test procedures prescribed these testers.

Let $\mathsf{T} = \{\mathsf{T}_\mathsf{M}, \mathsf{T}_\mathsf{N}\} \subset \mathcal{L}(\mathcal{H} \otimes \mathbb{C}^2)$ be binary 1-slot testers, with normalisation $\mathsf{T}_\mathsf{M} + \mathsf{T}_\mathsf{N} = \xi \otimes \mathbf{id}$, which describes our labeling experiments. Here, the process effect T_M corresponds to the decision that the observable has the labeling

$\{1 \mapsto M_1, 2 \mapsto M_2\}$ or is the observable M. Similarly, $\mathsf{T_M}$ corresponds to the decision that the observable has the labeling $\{1 \mapsto M_2, 2 \mapsto M_1\}$ or is the observable N. As such, our labeling experiment is characterised by probabilties of erroneous decisions: $p(\mathsf{N}|\mathsf{M})$ and $p(\mathsf{N}|\mathsf{M})$ which are given by $p(\mathsf{N}|\mathsf{M}) = \mathrm{Tr}\{\mathsf{T_N}\Phi_\mathsf{M}\}$ and $p(\mathsf{N}|\mathsf{M}) = \mathrm{Tr}\{\mathsf{T_M}\Phi_\mathsf{N}\}$. Now, the average error probability p_e is given by $\frac{1}{2}\mathrm{Tr}\{\mathsf{T_M}\Phi_\mathsf{N}\} + \frac{1}{2}\mathrm{Tr}\{\mathsf{T_N}\Phi_\mathsf{M}\}$. Minimisation of this quantity results in the following theorem.

Theorem 4.2.2. The single-shot optimal average error probability p_e^{opt} for labeling a binary observable $\mathsf{M} \in \mathbf{Obs}(\underline{2}, \mathcal{H})$, composed of effects M_1 and M_2, is given by

$$p_e^{\mathrm{opt}} = \frac{1}{2}(1 - \|M_1 - M_2\|_2).^{a} \tag{4.16}$$

[a] Where $\|X\|_2$ is the operator 2-norm, given by the maximal absolute eigenvalue of X.

Proof. We can start the proof by rewriting p_e using the normalisation of T,

$$\begin{aligned} p_e &= \frac{1}{2}\mathrm{Tr}\{\mathsf{T_M}\Phi_\mathsf{N}\} + \frac{1}{2}\mathrm{Tr}\{\mathsf{T_N}\Phi_\mathsf{M}\} && (4.17)\\ &= \frac{1}{2}\mathrm{Tr}\{\mathsf{T_M}\Phi_\mathsf{N}\} + \frac{1}{2}\mathrm{Tr}\{(\xi\otimes\mathbf{id} - \mathsf{T_M})\Phi_\mathsf{M}\} && (4.18)\\ &= \frac{1}{2}\mathrm{Tr}\{\mathsf{T_M}(\Phi_\mathsf{N} - \Phi_\mathsf{M}) + (\xi\otimes\mathbf{id})\Phi_\mathsf{M}\}. && (4.19) \end{aligned}$$

Substituting the expressions for the Choi operators, we have

$$\begin{aligned} p_e &= \frac{1}{2}\mathrm{Tr}\{\mathsf{T}_{12}(M_1^\top\otimes|0\rangle\langle 0| + M_2^\top\otimes|1\rangle\langle 1| - M_2^\top\otimes|0\rangle\langle 0| && (4.20)\\ &\quad - M_2^\top\otimes|1\rangle\langle 1|) + (\xi\otimes\mathbf{id})(M_2^\top\otimes|0\rangle\langle 0| + M_1^\top\otimes|1\rangle\langle 1|)\} && (4.21)\\ &= \frac{1}{2}\mathrm{Tr}\{\mathsf{T}_{12}((M_1^\top - M_2^\top)\otimes|0\rangle\langle 0| + (M_2^\top - M_1^\top)\otimes|1\rangle\langle 1|) + && (4.22)\\ &\quad \xi M_2^\top\otimes|0\rangle\langle 0| + \xi M_1^\top\otimes|1\rangle\langle 1|\}. && (4.23) \end{aligned}$$

Digressing, let us look at the trace of the summand term, $\xi M_2^\top\otimes|0\rangle\langle 0| + \xi M_1^\top\otimes|1\rangle\langle 1|$, and eventually incorporating the condition $M_1 + M_2 = \mathbf{id}$, we have

$$\begin{aligned} \mathrm{Tr}\left\{\xi M_2^\top\otimes|0\rangle\langle 0| + \xi M_1^\top\otimes|1\rangle\langle 1|\right\} &= \mathrm{Tr}\left\{\xi M_2^\top\right\}\otimes\langle 0|0\rangle + && (4.24)\\ &\qquad \mathrm{Tr}\left\{\xi M_1^\top\right\}\otimes\langle 1|1\rangle && (4.25)\\ &= \mathrm{Tr}\left\{\xi M_2^\top\right\}\otimes 1 + \mathrm{Tr}\left\{\xi M_1^\top\right\}\otimes 1 && (4.26)\\ &= \mathrm{Tr}\left\{\xi M_2^\top\right\} + \mathrm{Tr}\left\{\xi M_1^\top\right\} && (4.27)\\ &= \mathrm{Tr}\left\{\xi(M_2^\top + M_1^\top)\right\} && (4.28)\\ &= \mathrm{Tr}\left\{\xi(M_2 + M_1)^\top\right\} && (4.29)\\ &= \mathrm{Tr}\{\xi\mathbf{id}\} = 1. && (4.30) \end{aligned}$$

Now we have,

$$\begin{aligned} p_e &= \frac{1}{2}\mathrm{Tr}\left\{\mathsf{T}_{12}((M_1^\top - M_2^\top)\otimes|0\rangle\langle 0| + (M_2^\top - M_1^\top)\otimes|1\rangle\langle 1|)\right\} + \frac{1}{2} && (4.31)\\ &= \frac{1}{2}\mathrm{Tr}\left\{\mathsf{T}_{12}(M_1^\top - M_2^\top)\otimes(|0\rangle\langle 0| - |1\rangle\langle 1|)\right\} + \frac{1}{2}. && (4.32)\end{aligned}$$

Since the difference of two self-adjoint operators is a self-adjoint operator, we can have the spectral decomposition for the operator $(M_1^\top - M_2^\top) = \sum_{x=1}^d \mu_x |\omega_x\rangle\langle\omega_x|$, which furnishes the d real eigenvalues μ_x and the corrrespondingly associated eigenvectors $|\omega_x\rangle$. We can note that the two eigenvalues of the operator $(|0\rangle\langle 0| - |1\rangle\langle 1|)$ are 1 and -1. Consequently, the tensored operator $(M_1^\top - M_2^\top)\otimes(|0\rangle\langle 0| - |1\rangle\langle 1|)$ has $2d$ eigenvalues, $\mu_1,\ldots,\mu_d,-\mu_1,\ldots,-\mu_d$, including multiplicities. Moreover, we write this tensored operator, using its spectral decomposition, as

$$\begin{aligned} (M_1^\top - M_2^\top) &\otimes (|0\rangle\langle 0| - |1\rangle\langle 1|) && (4.33)\\ &= \left(\sum_x \mu_x |\omega_x\rangle\langle\omega_x|\right)\otimes(|0\rangle\langle 0| - |1\rangle\langle 1|) && (4.34)\\ &= \sum_x \mu_x |\omega_x\rangle\langle\omega_x|\otimes|0\rangle\langle 0| - \sum_x \mu_x |\omega_x\rangle\langle\omega_x|\otimes|1\rangle\langle 1| && (4.35)\\ &= \sum_x \mu_x |\omega_x\otimes 0\rangle\langle\omega_x\otimes 0| - \sum_x \mu_x |\omega_x\otimes 1\rangle\langle\omega_x\otimes 1|\,. && (4.36)\end{aligned}$$

Plugging in this form of the operator to the p_e, we have

$$p_e = \frac{1}{2} + \frac{1}{2}\sum_x \mu_x \langle\omega_x\otimes 0|\mathsf{T}_{12}|\omega_x\otimes 0\rangle - \frac{1}{2}\sum_x \mu_x \langle\omega_x\otimes 1|\mathsf{T}_{12}|\omega_x\otimes 1\rangle\,. \quad (4.37)$$

Since our task is to minimise this average error probability, we desire to suppress the positive summands, while maximising the negative summands, in tandem. Since T_{12} is a positive operator, the negativity (or positivity) of the terms $\mu_x \langle\omega_x\otimes j|\mathsf{T}_{12}|\omega_x\otimes j\rangle_{j=0,1}$ is solely dictated by the negativity (or positivity) of the values of μ_x. Moreover, we know that for each x that is associated with a nonzero valued μ_x, there are two summands, with one being positive and the other being negative. For example, let us say μ_1 is positive, then $\mu_1\langle\omega_1\otimes 0|\mathsf{T}_{12}|\omega_1\otimes 0\rangle$ is positive, while $-\mu_1\langle\omega_1\otimes 1|\mathsf{T}_{12}|\omega_1\otimes 1\rangle$ is negative. As such, we can rewrite the Eq.4.37 as having one positive summand and one negative summand by introducing the variables j_x and k_x, which are dependent on the positivity of the eigenvalues μ_x as follows; $j_x = 1$ and $k_x = 0$ when $\mu_x < 0$ and $j_x = 0$ and $k_x = 1$ when $\mu_x > 0$.

$$p_e = \frac{1}{2} + \underbrace{\frac{1}{2}\sum_x |\mu_x| \langle \omega_x \otimes j_x | \mathsf{T}_{12} | \omega_x \otimes j_x \rangle}_{\text{positive summand}} - \frac{1}{2}\sum_x |\mu_x| \langle \omega_x \otimes k_x | \mathsf{T}_{12} | \omega_x \otimes k_x \rangle \tag{4.38}$$

Now, essentially, we require such T_{12} those lead to the terms inside the positive summand being vanished: that is, $\mathsf{T}_{12} |\omega_x \otimes j_x\rangle = 0$ for all x. Thus, we have

$$p_e = \frac{1}{2} - \frac{1}{2}\sum_x |\mu_x| \langle \omega_x \otimes k_x | \mathsf{T}_{12} | \omega_x \otimes k_x \rangle . \tag{4.39}$$

Digressing, let us look at the positive numbers $q_x = \langle \omega_x \otimes k_x | \mathsf{T}_{12} | \omega_x \otimes k_x \rangle$. Due to the normalisation of T, we have $\mathsf{T}_{12} \leq \xi \otimes \mathbf{id}$, leading to

$$\begin{aligned} \langle \omega_x \otimes k_x | \mathsf{T}_{12} | \omega_x \otimes k_x \rangle &\leq \langle \omega_x \otimes k_x | \xi \otimes \mathbf{id} | \omega_x \otimes k_x \rangle && (4.40) \\ &= \langle \omega_x | \xi | \omega_x \rangle \langle k_x | k_x \rangle = \underbrace{\langle \omega_x | \xi | \omega_x \rangle}_{\text{probabilities}} . && (4.41) \end{aligned}$$

Noticing that $\langle \omega_x | \xi | \omega_x \rangle$ prescribe a probability distribution, we have $\sum_x q_x \leq \sum_x \langle \omega_x | \xi | \omega_x \rangle = 1$. We can infer that $q_x = \langle \omega_x \otimes k_x | \mathsf{T}_{12} | \omega_x \otimes k_x \rangle$ form, in general, a sub-normalised probability distribution. Now, with $p_e = \frac{1}{2}(1 - \sum_x |\mu_x| q_x)$, our goal to minimising p_e is translated into maximising the expression $\sum_x |\mu_x| q_x$ over all subnormalised probability distributions. This maximisation results in picking out the maximal $|\mu_x|$. Referring to the minimised (optimised) average error probability as p_e^{opt}, we have

$$\begin{aligned} p_e^{\text{opt}} &= \frac{1}{2}\left(1 - \max_{\{q_x\}} \left\{ \sum_x |\mu_x| q_x \right\}\right) \\ &= \frac{1}{2}(1 - \max_x \{|\mu_x|\}). && (4.42) \end{aligned}$$

Denoting $|\mu_m| = \max_x \{|\mu_x|\}$, the maximisation performed above is achieved when $\mathsf{T}_{12} = |\omega_m \otimes j_m\rangle\langle \omega_m \otimes j_m|$. Identifying that $|\mu_m|$ can be writen as the operator 2-norm of the operator $(M_1 - M_2)$, we can recast Equation 4.42 as the following, completing the proof.

$$p_e^{\text{opt}} = \frac{1}{2}(1 - \|M_1 - M_2\|_2). \tag{4.43}$$

□

Optimal labeling experiments

An optimal labeling experiment consists of preparing the state $|\omega_x\rangle$ which is the eigenvector of M_1, corresponding to the maximal or minimal eigenvalue λ_x, that maximises the value of $|\mu_x|$. When this single-shot procedure is carried out, we can note that the label corresponding to the effect M_1 is registered with probability $\lambda_x = \langle \omega_x | M_1 \omega_x \rangle$ and the one corresponding to M_2 with probability $\kappa_x = \langle \omega_x | M_2 \omega_x \rangle$. For the particular unlabeled observable M, if it is reckoned that $\lambda_x > \kappa_x$, then we can conclude that the registered label is associated with the effect M_1 and consequently the unregistered label with M_2. The averaged error associated with this procedure is characterised by the probability κ_x. This follows from Equation 4.42 and the fact that $\lambda_x = 1 - \kappa_x$,

$$\begin{aligned} p_e^{\text{opt}} &= \frac{1}{2}(1 - \max_x\{|\mu_x|\}) = \frac{1-(\lambda_x - \kappa_x)}{2} \\ &= \kappa_x. \end{aligned} \tag{4.44}$$

Similarly, if $\lambda_x < \kappa_x$, then we have $|\mu_x| = \kappa_x - \lambda_x$. In this scenario, we can associate the registered label with M_2 and the error of this association is characterised by λ_x. From the procedure we have sketched out till now, we can identify that in single-shot optimal minimum-error labeling experiments, we label the effect possessing the largest eigenvalue by implementing the unlabeled observable on the corresponding eigenvector state $|\omega_x\rangle$.

Coincidence with perfect labeling

From Equation 4.44, and the subsequent similar condition discussed after this equation above, we can read that the error probability is zero when either $\kappa_x = 0$ or $\lambda_x = 0$. This respectively means that either M_2 or M_1 should be rank-deficient, essentially implying that the observable itself should a rank-deficient one, in accordance with Theorem 4.2.1.

Operational distance related to labelability

In chapter 3, we discussed briefly the notion of operational distances between quantum objects, arising due to distinguishability tasks involving them. As a quick recollection, distinguishability of two equally probable states, ρ_1 and ρ_2, is expressed through the operational distance that is mediated by the operator trace norm $\|\rho_1 - \rho_2\|_1$. Inspired by this notion of operational distances, we can view the "labelability" of a binary observable, prescribed by the effect operators M_1 and M_2, as being expressed through the operational distance that is mediated by the operator 2-norm $\|M_1 - M_2\|_2$. In other words, we can read that the labelability of a given binary observable as this specific distance between the composing two effects.

For the case of two states which are identical to each other, that is, $\rho_1 = \rho_2$, as expected the operational distance between them is zero and consequently the minimum-error probability associated with their distinguishability is reckoned to be $\frac{1}{2}$. Even though this is the maximal value of error probability we can achieve in binary state distinguishability, across all possible pairs of states, associating an non-zero error to these cases when the states are identical can appear as pathological. A similar situation arises when we attempt to label observables. In contrast with the case of states, where we can have infinite number of pairs of identical states, when it comes to labeling, there exists a unique binary observable for each quantum system that gives rise to a similar pathological feature. This observable is the unbiased coin-toss observable, composed of effects $M_1 = \frac{\mathbf{id}}{2}, M_2 = \frac{\mathbf{id}}{2}$. It is evident that the operational distance between these two effects are zero, resulting in an error probabilty of half for labelability. Once you know that the two effects are identical to each other in prior, there is no need to conduct a labeling experiment or procedure.

Biased coin-toss observables

Binary coin-toss observables form the family of observables with effects $M_1 = q\mathbf{id}$ and $M_2 = (1-q)\mathbf{id}$ with $q \in [0,1]$. The unbiased case discussed in the previous passage corresponds to $q = \frac{1}{2}$. For a d dimensional system, the operator difference $M_1 - M_2$ has eigenvalue $(2q-1)$ with degeneracy d. Now, without loss of generality, we can consider the restricted family for $q \in [\frac{1}{2}, 1]$. Then the minimum error is given by,

$$p_e^{\mathrm{opt}} = 1 - q. \tag{4.45}$$

We can note that for $q = 1$, the error probability is zero and this corresponds to the trivial observables with effects $\mathbf{id}$ and O.

4.2.3 Imperfect unambiguous labeling

When a binary observable is known in advance that it seldom admits perfect labeling, that is, when we know both of the underlying non-identical effects, M_1 and M_2, are full rank operators, we turn to imperfect unambiguous labeling experiments. We require such labeling experiments not to admit erroneous decisions, but at the cost of admitting an inconclusive outcome, when the experimental round fails. As such, these experiments are described by 1-slot quantum testers composed of three effects $\mathsf{T} = \{\mathsf{T}_{\mathsf{M}}, \mathsf{T}_{\mathsf{N}}, \mathsf{T}_{?}\} \subset \mathcal{L}(\mathcal{H} \otimes \mathbb{C}^2)$, with normalisation $\mathsf{T}_{\mathsf{M}} + \mathsf{T}_{\mathsf{N}} + \mathsf{T}_{?} = \xi \otimes \mathbf{id}$. Here, the process effects T_{M} and T_{N} have the same identification as in the previous subsection, whereas $\mathsf{T}_{?}$ corresponds to the event when experimental failure occurs, leading to labeling inconclusion. These experiments reduce to the perfect labeling scenario when $\mathsf{T}_{?} = 0$, while we simultaneously require admittance of no erroneous decisions as well as no inconclusions.

We say that a binary observable $\mathsf{M} \in \mathbf{Obs}(\underline{2}, \mathcal{H})$ can be unambiguously labeled, in a single-shot, if there exists a non-trivial three process effect 1-slot quantum tester T, as described above, that satisfies the following no-error conditions,

$$\mathrm{Tr}\left\{\Phi_{\mathsf{M}}\mathsf{T}_{\mathsf{N}}\right\} = \mathrm{Tr}\left\{\Phi_{\mathsf{N}}\mathsf{T}_{\mathsf{M}}\right\} = 0. \tag{4.46}$$

Associated with the occurence of the inconclusive outcome, the performance of this experiment is quantified the average failure probability p_f, which is given by

$$p_f = \frac{1}{2}\mathrm{Tr}\left\{\mathsf{T}_?(\Phi_{\mathsf{M}} + \Phi_{\mathsf{N}})\right\}. \tag{4.47}$$

As such, our goal is to design non-trivial optimal testers T, satisfying Equation 4.46 with $\mathsf{T}_{\mathsf{M}} \neq O$ or $\mathsf{T}_{\mathsf{N}} \neq O$, which minimises the above average failure probability. With these discussed, we have the following no-go theorem on the admittance of imperfect unambiguous labeling.

> **Theorem 4.2.3.** A binary observable $\mathsf{M} \in \mathbf{Obs}(\underline{2}, \mathcal{H})$, that does not admit perfect labeling, cannot be imperfect-unambiguously labeled in a single shot.

Proof. Let $\mathsf{M} \in \mathbf{Obs}(\underline{2}, \mathcal{H})$ be the binary observable which does not admit perfect labeling. As such, by Theorem 4.2.1, the associated effects, M_1 and M_2 are full-rank operators and non-identical to each other. Now, plugging in the expressions for the Choi operators, Φ_{M} and Φ_{N}, into Equation 4.47, we have the average failure probability as

$$\begin{aligned}
p_f &= \frac{1}{2}\mathrm{Tr}\left\{\mathsf{T}_?(M_1^\top \otimes |0\rangle\langle 0| + M_2^\top \otimes |1\rangle\langle 1| + M_2^\top \otimes |0\rangle\langle 0| + M_1^\top \otimes |1\rangle\langle 1|)\right\} && (4.48)\\
&= \frac{1}{2}\mathrm{Tr}\left\{\mathsf{T}_?((M_1^\top + M_2^\top) \otimes |0\rangle\langle 0| + (M_2^\top + M_1^\top) \otimes |1\rangle\langle 1|)\right\} && (4.49)\\
&= \frac{1}{2}\mathrm{Tr}\left\{\mathsf{T}_?(\mathbf{id}_{\mathcal{H}} \otimes |0\rangle\langle 0| + \mathbf{id}_{\mathcal{H}} \otimes |1\rangle\langle 1|)\right\} && (4.50)\\
&= \frac{1}{2}\mathrm{Tr}\left\{\mathsf{T}_?(\mathbf{id}_{\mathcal{H}} \otimes (|0\rangle\langle 0| + |1\rangle\langle 1|))\right\} && (4.51)\\
&= \frac{1}{2}\mathrm{Tr}\left\{\mathsf{T}_?(\mathbf{id}_{\mathcal{H}} \otimes \mathbf{id}_{\mathbb{C}^2})\right\} && (4.52)\\
&= \frac{1}{2}\mathrm{Tr}\left\{\mathsf{T}_?(\mathbf{id}_{\mathcal{H}\otimes\mathbb{C}^2})\right\} && (4.53)\\
&= \frac{1}{2}\mathrm{Tr}\left\{\mathsf{T}_?\right\}. && (4.54)
\end{aligned}$$

Now, we want to minimise p_f over non-trivial testers, satisfying Equation 4.46. That is, to identify non-trivial solutions and then minimise p_f over these solutions.

For Equation 4.46 to be satisfied, for non-zero T_{M} and T_{N}, the operators Φ_{M} and Φ_{N} must have non-empty kernels. Let us consider the other ordering to write down Choi-Jamiołkowski operators, that is to have $\Lambda_{\mathsf{M}} \otimes \mathcal{I}(\Psi_+) = \tilde{\Phi_{\mathsf{M}}}$ and $\Lambda_{\mathsf{N}} \otimes \mathcal{I}(\Psi_+) = \tilde{\Phi_{\mathsf{N}}}$. These operators can be written as

$$\tilde{\Phi_{\mathsf{M}}} = \begin{pmatrix} M_1 & O \\ O & M_2 \end{pmatrix} \quad \text{and} \quad \tilde{\Phi_{\mathsf{N}}} = \begin{pmatrix} M_2 & O \\ O & M_1 \end{pmatrix}. \tag{4.55}$$

Since $\mathbf{rank}(A \otimes B) = \mathbf{rank}(B \otimes A)$, equivalently, we can look at these operators. Now, for these operators to have non-empty kernels, at least either M_1 or M_2 should be rank deficient. But in our case, since M_1 and M_2 are full rank-operators, they do not have non-empty kernels. Consequently, neither Φ_{M} or Φ_{N} has a non-empty kernel. From this, we can conclude that when the binary observable does not admit perfect labeling, neither does it allow imperfect unambiguous labeling. $\square$

If one of the non-identical effects, M_1 or M_2, is rank deficient, then by Theorem 4.2.1 the observable can be perfectly labeled. In such scenarios, we have $\mathsf{T}_? = O$, leading to $p_f = 0$, implying that for binary observables, unambiguous labeling (including perfect labeling as well as imperfect unambiguous labeling) reduces to just perfect labeling.

4.3 Non-binary observables

In the previous section, we investigated labeling tasks for binary observables. In this section, we proceed to investigate non-binary observables. As a quick recap, we refer to an observable $\mathsf{M} \in \mathbf{Obs}(\underline{n}, \mathcal{H})$ as being non-binary when $n \in [3, +\infty) \subset \mathbb{N}$, which is when it is composed of more than two effects. Analogous to labeling of binary observables being similar to binary distinguishability tasks, labeling tasks of non-binary are similar to distinguishability tasks of more than two non-identical states or any other quantum device.

Without loss of generality, let us assume that the given unlabeled non-binary observable $\mathsf{M} \in \mathbf{Obs}(\underline{n}, \mathcal{H})$ is composed of non-identical effects. As discussed earlier, to label the given observable is equivalent to distinguishing $n!$ possible observables, $\mathsf{M}_1, \mathsf{M}_2, \ldots, \mathsf{M}_{n!}$ with their respective, associated Choi-Jamiołkowski operators, $\Phi_1, \Phi_2, \ldots, \Phi_{n!}$.

4.3.1 Perfect labeling

When we want to perfectly label a non-binary observable with n non-identical effects, we need to simultaneously discriminate perfectly the corresponding $n!$

measurement channels pairwise. Attempting this task results in the following no-go theorem.

Theorem 4.3.1. A non-binary observable[a] $\mathsf{M} \in \mathbf{Obs}(\underline{n}, \mathcal{H})$, cannot be perfectly labeled in a single-shot.

[a]We stress that unbiased coin-toss observables are excluded.

Proof. As prescribed by Proposition 3.2.1, the required simultaneous perfect distinguishability conditions translate to requiring the existence of at least a tester T, composed of $n!$ process effects, with normalisation $\xi \otimes \mathbf{id}$, satisfying the following system of $\binom{n!}{2}$ equations.

$$\Phi_1(\xi \otimes \mathbf{id})\Phi_2 = O \tag{4.56}$$

$$\Phi_1(\xi \otimes \mathbf{id})\Phi_3 = O \tag{4.57}$$

$$\vdots \tag{4.58}$$

$$\Phi_{n!}(\xi \otimes \mathbf{id})\Phi_{(n-1)!} = O. \tag{4.59}$$

Let us analyse the first equation from this system, $\Phi_1(\xi \otimes \mathbf{id})\Phi_2 = O$, by plugging in the expressions for the Choi-Jamiołkowski operators appearing in it[4],

$$\Phi_1(\xi \otimes \mathbf{id})\Phi_2 = O \tag{4.60}$$

$$\left(\sum_j \mathsf{M}_1(j) \otimes |j\rangle\langle j|\right)(\xi \otimes \mathbf{id})\left(\sum_k \mathsf{M}_2(k) \otimes |k\rangle\langle k|\right) = O \tag{4.61}$$

$$\sum_j \sum_k \mathsf{M}_1(j)\xi\mathsf{M}_2(k) \otimes |j\rangle\langle j|\,|k\rangle\langle k| = O \tag{4.62}$$

$$\sum_j \mathsf{M}_1(j)\xi\mathsf{M}_2(j) \otimes |j\rangle\langle j| = O. \tag{4.63}$$

From this, we can read that $\mathsf{M}_1(j)\xi\mathsf{M}_2(j) = O$ should hold for all j. Similarly, the previous system of equations can be converted to one with $n \cdot \binom{n!}{2}$ equations. Based on the normalisation of the observable, it is always possible to carry out different cherry-pickings of some of these equations to get subsystems those when added together within individually gives a collection of equations $\mathsf{M}_1(j)\xi\mathbf{id} = O$. Summing this collection over j, we get $\mathbf{id}\xi\mathbf{id} = \xi = O$. This contradicts our assumption that there exists a tester with normalisation state ξ, thus concluding the proof. □

[4]We are omitting the transpositions of the effects since it does not affect the quality of the analysis.

4.3.2 Minimum-error labeling

Since we have understood that perfect labeling for a non-binary observable is not possible within the single-shot regime, we steer towards the study of minimum-error labeling of these observables. Prior to evaluating the optimal figure of merit for this task, we state and prove the following lemma which furnishes the optimal average success probability p_s for the entanglement-unassisted discrimination task of ensembles of states, those possess a specific structure. This lemma was presented and proven in [GZ21] as their Theorem 1. Since the lemma deals with the average success probability p_s, instead of the average error probability p_e, we incline towards choosing p_s and optimising it, for our labeling task as well. As discussed in the previous chapter, this should not cause any confusion since p_s and p_e are complementary figures of merit, especially when our experiements do not accommodate for any inconclusive outcomes.

Lemma 4.3.1. Let $\{(\varrho_k, \eta_k)\}_{k=1}^{N}$ be an ensemble of states with $\varrho_1, \ldots, \varrho_N$ being mutually commuting, thus sharing identical eigenprojectors $\{\Pi_k\}$ [a]. Then, the optimal average success probability p_s^{opt} correspodning to the discrimination of this ensemble of states without the assistance of entanglement is given by,

$$p_s^{\text{opt}} = \sum_k \operatorname{Tr}\{\Pi_k\} \max\{\operatorname{Tr}\{\eta_1 \varrho_1 \Pi_k\}, \ldots, \operatorname{Tr}\{\eta_N \varrho_N \Pi_k\}\}. \tag{4.64}$$

[a] Note that the commuting structure translates to the fact that all of these states share the same eigenprojectors Π_k.

Proof. As stated in the theorem, let $\varrho_1, \ldots, \varrho_N \in \mathbf{St}(\mathcal{H})$ be a finite collection of mutually commuting states, each respectively occuring with probabilities $\eta_1, \ldots, \eta_N$. Since these states commute with each other, they share the same system of eigenprojects, Π_k. Thus, these states can be spectral decomposed as $\rho_j = \sum_k w_{jk}\Pi_k$ where $w_{jk} = \operatorname{Tr}\{\varrho_j \Pi_k\}$ for all j. We can define an index set I_j that collects all those indices k for which $\eta_1 \operatorname{Tr}\{\varrho_j \Pi_k\}$ is maximised by ϱ_j. Then, the observable $\mathsf{F} \in \mathbf{Obs}(\underline{N}, \mathcal{H})$ distinguishing the collection of states is composed of effects $F_j = \sum_{k \in I_j} \Pi_k$, resulting in an associated success probability $p_s(\mathsf{F})$. Then, the optimal success probability is given by $p_s^{\text{opt}} = \max_{\mathsf{F}} p_s(\mathsf{F})$. This optimisation problem can be cast as an semi-definite program and the associated dual program reads [Bae13, BH13],

$$p_s^{\text{opt}} = \underbrace{\min_{K; K \geq \eta_j \varrho_j}}_{\text{for all } j} \operatorname{Tr}\{K\}. \tag{4.65}$$

We can define $x_k = \max\{\eta_1 \mathrm{Tr}\{\varrho_1 \Pi_k\}, \ldots, \eta_N \mathrm{Tr}\{\varrho_N \Pi_k\}\}$ and have $K = \sum_k x_k \Pi_k$. This choice is legit since $K \geq \eta_j \varrho_j$ for all j because $\sum_k \mathrm{Tr}\{\varrho_j \Pi_k\} \Pi_k$ guarantees it. We can conclude that this choice of K is optimal [Bae13, BH13] if we could identify at least one observable F that satisfies $\mathrm{Tr}\{F_j(K - \eta_j \varrho_j)\} = 0$, which is when F_j and $(K - \eta_j \varrho_j)$ have orthogonal supports for all j. Since we have $(K - \eta_j \varrho_j) = \sum_l (x_l - \eta_j w_{jl}) \Pi_l$, the operators $F_j = \sum_{l;\arg[\max_m w_{ml}]} \Pi_l$ are orthogonal to $(K - \eta_j \varrho_j)$ and we can also not that $\sum_j F_j = \mathbf{id}$. As such this observable leads to optimal success probability. □

Equipped with this lemma, we can now present our result as the following proposition.

Proposition 4.3.1. A non-binary observable $\mathsf{M} \in \mathbf{Obs}(\underline{n}, \mathcal{H})$, composed of effects $\{M_1, \ldots, M_n\}$, can be labeled without the assistance of entanglement, with an optimal average success probability,

$$p_s = \frac{1}{(n-1)!} \alpha_{\mathsf{M}}, \tag{4.66}$$

where $\alpha_{\mathsf{M}} = \max_{\rho, M_j} \mathrm{Tr}\{\rho M_j\}$ with ρ being the state of the system that is being measured.

Proof. We have already established that labeling corresponds to distinguishing between the possible $n!$ observables, $\mathsf{M}_1, \ldots, \mathsf{M}_{n!}$. When these observables are implemented on the system while it is in the state σ, this distinguishability translates itself to the one of states, $\rho_1, \ldots, \rho_{n!}$, each prepared by the respective measurement channels $\Lambda_{\mathsf{M}_1}, \ldots, \Lambda_{\mathsf{M}_{n!}}$. As such, they are given by $\rho_k = \sum_x \mathrm{Tr}\{\mathsf{M}_k(x)\sigma\} |x\rangle\langle x|$. We can observe that this collection of prepared states are mutually commuting. In the wake of the previous lemma, we can identify the projectors as $\{\Pi_m = |m\rangle\langle m|\}_{m=1}^n$. For labeling, since we assume that each of the labeled observables appear with equal chance, we have $\eta_k = \frac{1}{n!}$.

$$\begin{aligned} p_s \quad = \quad & \frac{1}{n!} \mathrm{Tr}\{|1\rangle\langle 1|\} \max\{\mathrm{Tr}\{\rho_1 |1\rangle\langle 1|\}, \ldots, \mathrm{Tr}\{\rho_{n!} |1\rangle\langle 1|\}\} + & (4.67) \\ & \frac{1}{n!} \mathrm{Tr}\{|2\rangle\langle 2|\} \max\{\mathrm{Tr}\{\rho_1 |2\rangle\langle 2|\}, \ldots, \mathrm{Tr}\{\rho_{n!} |2\rangle\langle 2|\}\} + \cdots + & (4.68) \\ & \frac{1}{n!} \mathrm{Tr}\{|n\rangle\langle n|\} \max\{\mathrm{Tr}\{\rho_1 |1\rangle\langle 1|\}, \ldots, \mathrm{Tr}\{\rho_{n!} |n\rangle\langle n|\}\}. & (4.69) \end{aligned}$$

We can now see that $\mathrm{Tr}\{\rho_k |m\rangle\langle m|\} = \mathrm{Tr}\{\mathsf{M}_k(m)\sigma\}$,

$$\begin{aligned} \mathrm{Tr}\{\rho_k |m\rangle\langle m|\} &= \mathrm{Tr}\left\{\left(\sum_x \mathrm{Tr}\{\mathsf{M}_k(x)\sigma\} |x\rangle\langle x|\right) |m\rangle\langle m|\right\} & (4.70) \\ &= \mathrm{Tr}\{\mathrm{Tr}\{\mathsf{M}_k(m)\sigma\} |m\rangle\langle m|\} & (4.71) \\ &= \mathrm{Tr}\{\mathsf{M}_k(m)\sigma\}. & (4.72) \end{aligned}$$

Since the associated observables have permuted label-effect assignments, the success probability takes the form,

$$
\begin{aligned}
p_s &= \frac{1}{n!}\max\{\mathrm{Tr}\{\sigma M_1\}, \mathrm{Tr}\{\sigma M_2\}, \ldots, \mathrm{Tr}\{\sigma M_n\}\} + &(4.73)\\
&= \frac{1}{n!}\max\{\mathrm{Tr}\{\sigma M_2\}, \mathrm{Tr}\{\sigma M_3\}, \ldots, \mathrm{Tr}\{\sigma M_1\}\} + \cdots + &(4.74)\\
&= \frac{1}{n!}\max\{\mathrm{Tr}\{\sigma M_n\}, \mathrm{Tr}\{\sigma M_1\}, \ldots, \mathrm{Tr}\{\sigma M_{n-1}\}\} &(4.75)
\end{aligned}
$$

Since each of the maximisations, appearing in the summands, are maximising over the identical collections of numbers, they are themselves identical to each other. This results in

$$
\begin{aligned}
p_s &= n\frac{1}{n!}\max\{\mathrm{Tr}\{\sigma M_1\}, \mathrm{Tr}\{\sigma M_2\}, \ldots, \mathrm{Tr}\{\sigma M_n\}\} &(4.76)\\
&= \frac{1}{(n-1)!}\max\{\mathrm{Tr}\{\sigma M_1\}, \mathrm{Tr}\{\sigma M_2\}, \ldots, \mathrm{Tr}\{\sigma M_n\}\} &(4.77)\\
&= \frac{1}{(n-1)!}\max_{M_k}\{\mathrm{Tr}\{\sigma M_k\}\}. &(4.78)
\end{aligned}
$$

In our case, we can observe that the state σ is a variable as well. Thus, we find the optimal average success probability by optimising the above p_s over all possible states,

$$
\begin{aligned}
p_s^{\mathrm{opt}} &= \frac{1}{(n-1)!}\max_{M_k,\sigma}\{\mathrm{Tr}\{\sigma M_k\}\}\\
&= \frac{1}{(n-1)!}\alpha_{\mathsf{M}}. &(4.79)
\end{aligned}
$$

Here, we have defined α_{M} to be $\max_{M_k,\sigma}\{\mathrm{Tr}\{\sigma M_k\}\}$, thus completing the proof. □

The optimal success probability p_s^{opt} we arrived at is for the restricted strategies where entanglement-assistance is not permitted. This renders us curious whether entanglement-assisted protocols improve this p_s^{opt}. On top of that, whether adaptive protocols improve it is also the next natural question to ask. Both of these can be simultaneously addressed when we investigate the problem using testers. Consequently, we find that neither entanglement-assistance nor adaptivity of protocols result in the improvement of the figure of merit. This is captured by the following theorem.

Theorem 4.3.2. The optimal success probability p_s^{opt} achieved without the assitance of entanglement, given by Equation 4.66, is optimal over all possible strategies. In other words, there does not exist an entanglement-assisted strategy that can improve p_s.

Proof. To optimise the labeling procedure for our observable, over all possible strategies including entanglement-assisted ones, is to optimise the figure of merit over general 1-slot testers T. That is, we need to maximise the average success probability $p_s(\mathsf{T}) = \frac{1}{n!}\left\{\sum_k \mathrm{Tr}\left\{\mathsf{T}_k\Phi_k\right\}\right\}$to arrive at the optimal one,

$$p_s^{\mathrm{opt}} = \frac{1}{n!}\max_{\mathsf{T}}\left\{\sum_k \mathrm{Tr}\left\{\mathsf{T}_k\Phi_k\right\}\right\}. \tag{4.80}$$

We have seen that, without loss of generality, the set of testers can be restricted to those of the form $\mathsf{T}_c = \sum_i H_i^{(c)} \otimes |i\rangle\langle i|$ with $\sum_c H_i^{(c)} = \xi$ for all i, where ξ is the normalisation state of the tester. Let us evaluate p_s^{opt} for the case of the observable $\mathsf{M} \in \mathbf{Obs}(\underline{3},\mathcal{H})$, composed of non-identical effects M_1, M_2, and M_3. Then, there are six non-identical observables and their corresponding Choi-Jamiołkowski operators, Φ_j. The corresponding tester is composed of six process effects, $\mathsf{T}_c = \sum_{i=1}^3 H_i^{(c)} \otimes |i\rangle\langle i|$. Plugging in the expressions for these operators, we arrive at the six success probabilities as,

$$\begin{aligned}
\mathrm{Tr}\left\{\mathsf{T}_1\Phi_1\right\} &= \mathrm{Tr}\left\{M_1H_1^{(1)}\right\} + \mathrm{Tr}\left\{M_2H_2^{(1)}\right\} + \mathrm{Tr}\left\{M_3H_3^{(1)}\right\} \\
\mathrm{Tr}\left\{\mathsf{T}_2\Phi_2\right\} &= \mathrm{Tr}\left\{M_1H_1^{(2)}\right\} + \mathrm{Tr}\left\{M_2H_2^{(2)}\right\} + \mathrm{Tr}\left\{M_3H_3^{(2)}\right\} \\
&\vdots \\
\mathrm{Tr}\left\{\mathsf{T}_6\Phi_6\right\} &= \mathrm{Tr}\left\{M_3H_1^{(6)}\right\} + \mathrm{Tr}\left\{M_2H_2^{(6)}\right\} + \mathrm{Tr}\left\{M_1H_3^{(6)}\right\}.
\end{aligned} \tag{4.81}$$

Plugging these probabilities into $p_s(\mathsf{T})$ and introducing the notation $H_x^{\alpha,\beta} = H_x^{(\alpha)} + H_x^{(\beta)}$ for convenience, we arrive at the following success probability after some regrouping of the terms involved.

$$\begin{aligned}
p_s(\mathsf{T}) &= \frac{1}{3!}\mathrm{Tr}\left\{M_1H_1^{12} + M_2H_1^{34} + M_3H_1^{56}\right\} \\
&\quad+\frac{1}{3!}\mathrm{Tr}\left\{M_1H_2^{35} + M_2H_2^{16} + M_3H_2^{24}\right\} \\
&\quad+\frac{1}{3!}\mathrm{Tr}\left\{M_1H_3^{46} + M_2H_3^{25} + M_3H_3^{13}\right\}.
\end{aligned} \tag{4.82}$$

Let us consider the first summand in the above equation, with the factor of $\frac{1}{3!}$ removed, as $\Lambda = \mathrm{Tr}\left\{M_1H_1^{12} + M_2H_1^{34} + M_3H_1^{56}\right\}$. With our prior knowledge that $H_1^{12} + H_1^{34} + H_1^{56} = \xi$, without loss of generality, we can assume that $H_1^{12} = w_1\sigma_1, H_1^{34} = w_2\sigma_2,$ and $H_1^{56} = w_3\sigma_3$, where w_i are probabilities adding up to one and σ_i are some states. Then, we have

$$\Lambda = w_1\mathrm{Tr}\left\{M_1\sigma_1\right\} + w_2\mathrm{Tr}\left\{M_2\sigma_2\right\} + w_3\mathrm{Tr}\left\{M_3\sigma_3\right\} \tag{4.83}$$

$$\leq w_1\max_{\sigma_1}\mathrm{Tr}\left\{M_1\sigma_1\right\} + w_2\max_{\sigma_2}\mathrm{Tr}\left\{M_2\sigma_2\right\} + w_3\max_{\sigma_3}\mathrm{Tr}\left\{M_3\sigma_3\right\} \tag{4.84}$$

$$\leq \max\{\max_{\sigma_1}\mathrm{Tr}\left\{M_1\sigma_1\right\} + \max_{\sigma_2}\mathrm{Tr}\left\{M_2\sigma_2\right\} + \max_{\sigma_3}\mathrm{Tr}\left\{M_3\sigma_3\right\}\} \tag{4.85}$$

$$= \max_{\sigma.M_i}\mathrm{Tr}\left\{M_i\sigma\right\} = \alpha_{\mathsf{M}}.$$

We find that Λ is upper bounded by α_{M}. Carrying out similar calculations for the rest of the terms appearing in $p_s(\mathsf{T})$ we find that they are also individually bounded above by α_{M}. As such, we can bound the success probability from above as follows,

$$\begin{aligned} p_s(\mathsf{T}) &\leq \frac{1}{3!}\alpha_{\mathsf{M}} + \frac{1}{3!}\alpha_{\mathsf{M}} + \frac{1}{3!}\alpha_{\mathsf{M}} \\ &= \frac{1}{2!}\alpha_{\mathsf{M}} = \frac{1}{(3-1)}\alpha_{\mathsf{M}}. \end{aligned} \tag{4.86}$$

We can extend this analysis to non-binary observables composed of more than three effects. Consequently, for a non-binary observable with n effects, the success probability is upper bounded as follows,

$$p_s(\mathsf{T}) \leq \frac{1}{(n-1)!}\alpha_{\mathsf{M}}. \tag{4.87}$$

Comparing this result with Equation 4.79, we can learn that the upper bound on this probability of success is achieved by optimal labeling experiments which are not assisted by entanglement. As such, we can conclude that entanglement assistance does not improve labelability of non-binary observables. □

A comment on binary observables

We can note that even though we have presented the result for non-binary observables, this result is also valid for binary observables. This reveals the fact that optimal success probability achieved in minimum-error labeling, $p_s^{\text{opt}} = \frac{1}{2}(1 + \|M_1 - M_2\|_2)$, can also be achieved without the assistance of entanglement.

von Neumann observables

The d number of effects of a von Neumann observable $\mathsf{V} \in \mathbf{Obs}(\underline{d}, \mathcal{H})$, $V_1, \ldots, V_d$, are all rank-1 projectors. We can evaluate the optimal success probability, $p_s^{\text{opt}} = \frac{1}{(d-1)!} \max_{V_k,\sigma}\{\mathrm{Tr}\left\{\sigma V_1\right\}, \ldots, \mathrm{Tr}\left\{\sigma V_d\right\}\}$. By choosing the states $\sigma = V_1$, we find that the maximisation is achieved, yielding the value of one; In fact, we can choose any of the involved effects. Thus, here $\alpha_{\mathsf{V}} = 1$, resulting in the optimal success probability,

$$p_s^{\text{opt}} = \frac{1}{(d-1)!}. \tag{4.88}$$

We find that optimal labelability of von Neumann observables is a sole function of the system and that it decreases with increasing dimension.

Trine observables

We can look at the following choice of trine observable $\mathsf{M} \in \mathbf{Obs}(\underline{3}, \mathbb{C}^2)$,

$$M_1 = \frac{2}{3} |0\rangle\langle 0| , \quad M_2 = \frac{2}{3} |v_+\rangle\langle v_+| , \quad M_3 = \frac{2}{3} |v_-\rangle\langle v_-| . \tag{4.89}$$

with $|v_\pm\rangle = \frac{1}{2} |0\rangle \pm \frac{\sqrt{3}}{2} |1\rangle$. We can see that $\alpha_{\mathsf{M}} = \frac{2}{3}$ which can be achieved by using any of the states $|0\rangle$, $|v_+\rangle$ or $|v_-\rangle$. This enables us to write the optimal success probability as

$$p_s^{\text{opt}} = \frac{1}{2} \cdot \frac{2}{3} = \frac{1}{3}. \tag{4.90}$$

Coin-toss observables

Given a coin-toss observable $\mathsf{C} \in \mathbf{Obs}(\underline{n}, \mathcal{H})$, with effects $C_1 = q_1 \mathbf{id}, \ldots, C_n = q_n \mathbf{id}$, we need to evaluate $\alpha_{\mathsf{C}} = \max_{\sigma, C_i} \{\mathrm{Tr}\{C_i \sigma\}\}$,

$$\begin{aligned} \alpha_{\mathsf{C}} &= \max_\sigma \{\mathrm{Tr}\{q_i \mathbf{id} \sigma\}\} = \max_\sigma \{q_i \mathrm{Tr}\{\sigma\}\} \\ &= \max\{q_i\}. \end{aligned} \tag{4.91}$$

Denoting $q_{\max}$ as the largest of the probabilities q_i, we can write down the optimal success probability as,

$$p_s^{\text{opt}} = \frac{1}{(n-1)!} q_{\max}. \tag{4.92}$$

4.3.3 Imperfect unambiguous labeling

We have seen that a non-binary observable does not admit perfect labeling in a single shot. Alongside, consider this specific scenario where we initiate to consider the multiple-shot case for labeling. Assume that we have a von Neumann observable $\mathsf{N} \in \mathbf{Obs}(\underline{3}, \mathbb{C}^3)$, composed of the effects $|1\rangle\langle 1| , |2\rangle\langle 2|$,and $|3\rangle\langle 3|$. As we have discussed, this observable does not admit perfect labeling in a single shot. Moreover, when we consider two shots of the observable, we can perfectly label this observable; One procedure is by measuring the state $|1\rangle$ in the first shot and then $|2\rangle$ in the second shot, the corresponding recorded outcome labels correspond to the respective effects $|1\rangle\langle 1|$ and $|2\rangle\langle 2|$ and the third unrecorded outcome label corresponds to the effect $|3\rangle\langle 3|$. In conclusion, this is a scenario where the observable does not admit perfect labeling in a single shot but does in at least two shots. Such scenarios prompt us to ask the question, among other similar questions, that whether an observable, which is perfectly labelable in at least two shots, admit imperfect unambiguous labeling in a single shot by incorporating an inconclusive outcome.

We say that a non-binary observable $\mathsf{M} \in \mathbf{Obs}(\underline{n}, \mathcal{H})$, composed of non-identical effects, admit unambiguous labeling if there exists at least one 1-slot quantum tester, having $(n!+1)$ process effects $\{\mathsf{T}_1, \ldots, \mathsf{T}_{n!}, \mathsf{T}_?\}$ and normalisation $\sum_j \mathsf{T}_j + \mathsf{T}_? = \xi \otimes \mathbf{id}$, such that the following no-error conditions are met for at least one non-trivial T_j,

$$\mathrm{Tr}\left\{\Phi_x \mathsf{T}_j\right\} = 0 \quad \text{for all } x, j \text{ and } j \neq x, ? \tag{4.93}$$

If there are more than one such testers, then our task is to find those which minimise the failure probability $p_f = \mathrm{Tr}\left\{\mathsf{T}_? \left(\sum_j \Phi_j\right)\right\}$. Subsequent investigation of this labeling results in the following no-go theorem.

Theorem 4.3.3. A non-binary observable $\mathsf{M} \in \mathbf{Obs}(\underline{n}, \mathcal{H})$ does not admit imperfect unambiguous labeling in a single-shot.

Proof. Let us the study the case of $\mathsf{M} \in \mathbf{Obs}(\underline{3}, \mathcal{H})$ with non-identical effects. From the no-error conditions, we have

$$\mathrm{Tr}\left\{\Phi_1 \mathsf{T}_2\right\} = \mathrm{Tr}\left\{\Phi_1 \mathsf{T}_3\right\} = \cdots = \mathrm{Tr}\left\{\Phi_1 \mathsf{T}_6\right\} = 0. \tag{4.94}$$

Substituting, the expressions for the Choi-Jamiłkowski operators and the process effects, we arrive at

$$\mathrm{Tr}\left\{H_1^{(2)} M_1\right\} = \mathrm{Tr}\left\{H_1^{(3)} M_1\right\} = \cdots = \mathrm{Tr}\left\{H_1^{(6)} M_1\right\} = 0. \tag{4.95}$$

Then, from the next batch of no-error conditions, we can pick out $\mathrm{Tr}\left\{\Phi_2 \mathsf{T}_1\right\} = 0$, from which we can have $\mathrm{Tr}\left\{H_1^{(1)} M_1\right\} = 0$. Adding this equation and the string of equations appearing above, we have

$$\begin{aligned} 0 &= \sum_c \mathrm{Tr}\left\{H_1^{(c)} M_1\right\} = \mathrm{Tr}\left\{\left(\sum_c H_1^{(c)}\right) M_1\right\} && (4.96) \\ &= \mathrm{Tr}\left\{\xi M_1\right\}. && (4.97) \end{aligned}$$

Carrying out similar analyses, we arrive at $\mathrm{Tr}\left\{\xi M_2\right\} = 0$ and $\mathrm{Tr}\left\{\xi M_3\right\} = 0$. Consequently, adding up these equations, we have $\mathrm{Tr}\left\{\rho M_1\right\} + \mathrm{Tr}\left\{\xi M_2\right\} + \mathrm{Tr}\left\{\xi M_3\right\} = \mathrm{Tr}\left\{\xi\right\} = 0$. This implies that there does not exist a valid normalisation state ξ and consequently a 1-slot tester that can imperfect unambiguously label the observable. This proof can be extended to any non-binary observable $\mathsf{M} \in \mathbf{Obs}(\underline{n}, \mathcal{H})$. □

***Observation* 4.3.1.** It should be noted that the above proof covers the whole scenario of unambiguous labeling, rather than just imperfect unambiguous labeling. As such, the result that non-binary observables does not admit perfect labeling in a single-shot, expressed as Theorem 4.3.1, can be seen as a corollary to the theorem we just presented. The two theorems can be read together as "*A non-binary observable does not admit unambiguous labeling in a single shot*".

4.4 Partial labeling and Antilabeling

4.4.1 Partial labeling of observables

Consider the following effects of an observable $\mathsf{M} \in \mathbf{Obs}(\underline{3}, \mathbb{C}^2)$,

$$M_1 = \begin{pmatrix} 0 & 0 \\ 0 & 0.7 \end{pmatrix}, \; M_2 = \begin{pmatrix} 1 & 0 \\ 0 & 0.2 \end{pmatrix}, \; M_3 = \begin{pmatrix} 0 & 0 \\ 0 & 0.1 \end{pmatrix}. \tag{4.98}$$

As we have seen, it is not possible to perfectly label this non-binary observable in a single-shot. Regardless, we can notice that one of the effects, M_2, has an eigenvalue 1. As such, let us pick the state $|0\rangle$, from the subspace spanned by the eigenvector corresponding this eigenvalue 1. Measuring the observable on this state will always register the outcome label corresponding to the effect M_2. Since we are partially labeling the observable by identifying a single labeling, we refer to this task as *partial labeling*. Now, consider the effects of the observable $\mathsf{N} \in \mathbf{Obs}(\underline{4}, \mathbb{C}^2)$,

$$M_1 = \begin{pmatrix} 0.4 & 0 \\ 0 & 0.2 \end{pmatrix}, \; M_2 = \begin{pmatrix} 0.6 & 0 \\ 0 & 0.3 \end{pmatrix}, \; M_3 = \begin{pmatrix} 0 & 0 \\ 0 & 0.4 \end{pmatrix}, \; M_4 = \begin{pmatrix} 0 & 0 \\ 0 & 0.1 \end{pmatrix}. \tag{4.99}$$

Observing that $M_1 + M_2$ has an eigenvalue 1, we can perform the identical process of measuring this observable on the state $|0\rangle$. From the recorded outcome label, we can infer, the non-trivial information regarding the label, that it either corresponds to the effect M_1 or to the effect M_2. On the other hand, we can observe that due to the normalisation of the observable, $M_1 + M_2 + M_3 + M_4$ also has an eigenvalue 1, corresponding to $|1\rangle$. Subsequent measurement on the state $|1\rangle$ reveals that the recorded label corresponds to any of the four effects, which is nothing but trivial information.

***Observation* 4.4.1.** If the observable M, composed of effects $M_1, \ldots, M_n$, is full rank, then this scheme of partial labeling is not possible to implement, as we cannot extract any non-trivial information. As such, when the observable is necessarily rank-deficient , there exists at least one strict subset of effects $\{M_a\}_a \subset \{M_1, \ldots, M_n\}$ such that there exists a state $|\varphi\rangle$ such that $(\sum_a M_a)\,|\varphi\rangle = |\varphi\rangle$.

***Observation* 4.4.2.** We can observe that for a binary observable, if it admits partial labeling, then it is equivalent to perfect labeling. So, essentially when we refer to partial labeling of observables, we refer to the relevant case of non-binary observables.

After looking at the above two cases and drawing few observations regarding the structures present, we can formalise the notion, and associated procedure, as follows.

> Given a non-binary, rank-deficient observable $\mathsf{M} \in \mathbf{Obs}(\underline{n}, \mathcal{H})$, composed of the effects $\{M_1, \ldots, M_n\}$, there exists at least a strict subset of m effects, $X_\varphi \subset \{M_1, \ldots, M_n\}$, such that the operator defined as $M_\varphi = \sum_{M_j \in X_\varphi} M_j$ satisfies the relation $M_\varphi |\varphi\rangle = |\varphi\rangle$.

The presented notion of partial labeling will prove itself to have more potential, for specifically structured observables, in the multiple-shot regime.

4.5 Anti-labeling

We have understood that having eigenvalue 1 determines the labelability of effects and, consequently, of observables. This has led us to formulate the tasks of partial labeling, as discussed above. Complementary to this observation, we can identify the notion of "anti-labelability" of effects and observables; These schemes are motivated from antidistinguishability tasks [BJOP14, CFS02, HK18] of states, where we are tasked to exclude states, in constrast to identifying them through distinguishability tasks. As an auxiliary exercise, consider the following observble,

$$M_1 = \begin{pmatrix} 0 & 0 \\ 0 & 0.7 \end{pmatrix}, \; M_2 = \begin{pmatrix} 0.8 & 0 \\ 0 & 0.2 \end{pmatrix}, \; M_3 = \begin{pmatrix} 0.2 & 0 \\ 0 & 0.1 \end{pmatrix}. \tag{4.100}$$

We can notice that M_2 and M_3 are full-rank operators, while M_1 being rank-deficient with $M_1 |0\rangle = 0$. As such, when we measure this observable on the state $|0\rangle$, we can conclude that the corresponding recorded label does not correspond to the effect M_1. We can now say that we have *anti-labeled* the effect M_1. Now, consider the effects of the observable $\mathsf{N} \in \mathbf{Obs}(\underline{4}, \mathbb{C}^2)$,

$$M_1 = \begin{pmatrix} 0.4 & 0 \\ 0 & 0.2 \end{pmatrix}, \; M_2 = \begin{pmatrix} 0.6 & 0 \\ 0 & 0.3 \end{pmatrix}, \; M_3 = \begin{pmatrix} 0 & 0 \\ 0 & 0.4 \end{pmatrix}, \; M_4 = \begin{pmatrix} 0 & 0 \\ 0 & 0.1 \end{pmatrix}. \tag{4.101}$$

We can now notice that both M_3 and M_4 are rank-deficient as well as share the same null space, with $(M_3 + M_4)|0\rangle = 0$. As such, we perform a similar measurement, measuring this observable on the state $|0\rangle$. From this single-shot measurement, we can conclude that the recorded outcome label does not correspond to either M_3 or M_4.

***Observation* 4.5.1.** As with the case of partial labeling, when the observable is full-rank, anti-labeling is not admitted. Moreover, every rank-deficient observable admits anti-labeling to varying capacities, depending on the eigenvalue structure of the rank-deficient effects involved. This is anyway relevant only in the multiple-shot regime of this task.

***Observation* 4.5.2.** Again, as with the case of partial labeling, anti-labeling of a binary observable, if it admits this task, is equivalent to perfect labeling. Consequently, we are interested with anti-labeling tasks concerning non-binary observables.

Enabled by this discussion, we can formalise this notion as follows.

A non-binary observable $\mathsf{M} \in \mathbf{Obs}(\underline{n}, \mathcal{H})$, composed of effects $\{M_1, \dots M_n\}$, admits anti-labeling if it is a rank-deficient observable. That is, there exists at least one subset of m effects, $Y_\varphi \subset \{M_1, \dots M_n\}$, such that the operator defined as $\tilde{M}_\varphi = \sum_{M_j \in Y_\varphi} M_j$ satisfies the relation $\tilde{M}_\varphi |\varphi\rangle = 0$. We can anti-label the set Y_φ by measuring the observable on the state $|\varphi\rangle$.

Chapter 5

Multiple-shot labeling

In the previous chapter, we introduced and mathematically formulated the task of labeling unlabeled observables and investigated this task within the single-shot regime. Even though we have not explicitly stressed enough in that chapter, single-shot regime addresses the relevant question of what can we learn, as restricted by the theory, about the unlabeled observable regarding its labeling when a single use or query of the observable is available. But, in general, the measurement device implementing the unlabeled observable might be available for more than a single use or query. In such scenarios, we can ask the following non-exhaustive list of questions:

– The skeletal question being "Whether we can learn more about the labelability of an unlabeled observable when we have access to more shots than what we can do with a single shot?"

– Can we perfectly label with multiple shots those observables which cannot be perfectly labeled in a single shot? Same question can be asked for imperfect unambiguous labeling experiments.

– Are there any specific minimal numbers of shots which can achieve perfect labelability?

As such, the investigations in this chapter are inspired by these questions as well as the central themes of labelability which we dealt with in the previous chapter. Following the structure of the previous chapter, we investigate binary observables and then proceed to the more general class of non-binary observables. We adopt the same notations established earlier and avoid repetition by not attempting to establish them again in this chapter. Then we investigate the tasks of partial labeling and anti-labeling separately.

Schemes for multiple-shot labeling

As discussed in subsection 3.3.2, labelability schemes can be either parallel or adaptive in general. On top of the potential advantage due to accessing more number of copies or queries of the unlabeled observable, parallel and adaptive are, in general, both entanglement-assisted schemes as well. On top of assistance due to entanglement, adpative schemes also possess quantum memories between each implementation of the measurement device, those which could also result in enhanced labelability.

5.1 Binary observables

5.1.1 Perfect labeling

From Theorem 4.2.1, we learnt that full-rank binary observables, composed of non-identical effects, do not admit perfect labeling in a single-shot. Here, we investigate whether such observables can be perfectly labeled if more uses or queries of observables are available. First, we start this investigation with parallel schemes which results in the upcoming propostion. When N copies of the unlabeled binary observable M is available with $[\mathsf{M}] = \{\mathsf{M}, \mathsf{N}\}$, we say that the observable is perfectly labelable using parallel schemes if there exists a 1-slot quantum tester, with normalisation $\xi \otimes \mathbf{id}$, that satisfies the following condition,

$$\Phi_{\mathsf{M}}^{\otimes N}(\xi \otimes \mathbf{id})\Phi_{\mathsf{N}}^{\otimes N} = O. \tag{5.1}$$

Proposition 5.1.1. If a binary observable $\mathsf{M} \in \mathbf{Obs}(\underline{2}, \mathbb{C}^2)$ does not admit perfect labeling in a single shot, then it does not admit perfect labeling in any finite number of shots, when implemented in parallel.

Proof. Assume that the unlabeled observable M does not admit itself to be perfectly labeled in a single shot and consequently composed of non-identical, full-rank effects. Let us analyse this task for the case when two copies of the observable are available. As such $\Phi_{\mathsf{M}}^{\otimes 2}(\xi \otimes \mathbf{id})\Phi_{\mathsf{N}}^{\otimes 2} = O$ needs to be satisfied. Plugging in the Choi-Jamiłkowski operators, we have

$$\begin{aligned}
&\sum_{j,k,x,y} \big(\mathsf{M}(j)^\top \otimes \mathsf{M}(k)^\top \otimes |j\rangle\langle j| \otimes |k\rangle\langle k|\big) [\xi \otimes \mathbf{id}] \big(\mathsf{N}(x)^\top \otimes \mathsf{N}(y)^\top \otimes |x\rangle\langle x| \otimes |y\rangle\langle y|\big) \\
&\quad = \sum_{x,y} \big((\mathsf{M}(j)^\top \otimes \mathsf{M}(k)^\top)[\xi](\mathsf{N}(x)^\top \otimes \mathsf{N}(y)^\top)\big) \otimes |j\rangle\langle j|\,|x\rangle\langle x| \otimes |k\rangle\langle k|\,|y\rangle\langle y| \\
&\qquad = \sum_{x,y} (\mathsf{M}(x)^\top \otimes \mathsf{M}(y)^\top)[\xi](\mathsf{N}(x)^\top \otimes \mathsf{N}(y)^\top) \otimes |x\rangle\langle x| \otimes |y\rangle\langle y| = O
\end{aligned} \tag{5.2}$$

Since, each of the summands in the above sum is a positive semidefinite operator, for their sum to be equal to the zero operator O each of them should separately be equal to the zero operator. That is, $((\mathsf{M}(x)^\top \otimes \mathsf{M}(y)^\top)\xi(\mathsf{N}(x)^\top \otimes \mathsf{N}(y)^\top)) \otimes |x\rangle\langle x| \otimes |y\rangle\langle y| = O$ for all x and y. Moreover, since the operators $|x\rangle\langle x| \otimes |y\rangle\langle y|$ are positive, for all x and y, it follows that the following equalities should be satisfied for all x and y.

$$(\mathsf{M}(x)^\top \otimes \mathsf{M}(y)^\top)[\xi](\mathsf{N}(x)^\top \otimes \mathsf{N}(y)^\top) = O \tag{5.3}$$

Now, since the observable M does not admit perfect labeling, the effects, $\mathsf{M}(1) = \mathsf{N}(2)$ and $\mathsf{M}(2) = \mathsf{N}(1)$, are full rank operators, as well as their transposed effects. This ensures that inverses of these operators exist as well and enabling us to sandwich the above equation, left with $((\mathsf{M}(x)^\top)^{-1} \otimes (\mathsf{M}(y)^\top)^{-1})$ and right with $((\mathsf{N}(x)^\top)^{-1} \otimes (\mathsf{N}(y)^\top)^{-1})$. This leads to

$$\xi = O. \tag{5.4}$$

Since $\xi = O \notin \mathbf{St}(\mathcal{H} \otimes \mathcal{H})$, we can assert that there does not exist a binary 1-slot tester, into which the two copies of the observable are inserted in parallel, that result in the perfect labeling of the observable. This completes the proof for the case of parallel use of two copies of the observable, but can be generalised straightforwardly to the case of any finite N uses. $\square$

Entanglement-assisted parallel schemes

Having explored the case of parallel schemes and finding that they seldom aid in achieving perfect labelability for those specific observables, we can learn that entanglement-assistance as well as multiple copy parallel schemes do not enhance labelability to perfection. At this point, we might have lost hope in entanglement-assistance but quantum memories could still aid in achieving perfection. This leads us to the following investigation with of adaptive schemes.

We can assert that the binary observable, which we were considering till now, can be perfectly labeled with N queries using adaptive schemes if there exists a N-slot quantum tester, with normalisation $\Xi^{(N)} \otimes \mathbf{id}$, that satisfies the following condition,

$$\Phi_{\mathsf{M}}^{\otimes N}(\Xi^{(N)} \otimes \mathbf{id})\Phi_{\mathsf{N}}^{\otimes N} = O. \tag{5.5}$$

Investigating this condition, we end up with the following proposition.

Proposition 5.1.2. If a binary observable $\mathsf{M} \in \mathbf{Obs}(\underline{2}, \mathbb{C}^2)$ does not admit perfect labeling in a single shot, then it does not admit perfect labeling in any finite number of shots, when implemented using adaptive strategies.

Proof. Similar to the previous proof, let at us look at the case when we have access to two copies of the observable. As such, we need find a binary 2-slot quantum tester $\mathsf{T}^{(2)}$, with process effects $\{\mathsf{T}_1^{(2)}, \mathsf{T}_2^{(2)}\} \subset \mathcal{L}(\mathcal{H}_0 \otimes \mathcal{H}_2 \otimes \mathcal{H}_1 \otimes \mathcal{H}_3)$ and normalisation conditions $\sum_k \mathsf{T}_k^{(2)} = \Xi^{(2)} \otimes \mathbf{id}_3$, $\mathrm{Tr}_2\{\Xi^{(2)}\} = \xi \otimes \mathbf{id}_1$ and $\mathrm{Tr}\,\{\xi\} = 1$, satisfying the following perfect labelability condition.

$$\Phi_{\mathsf{M}}^{\otimes 2}[\Xi^{(2)} \otimes \mathbf{id}_3]\Phi_{\mathsf{N}}^{\otimes 2} = O. \tag{5.6}$$

Plugging in the expressions for the Choi-Jamiołkowski operators, we have

$$\begin{aligned}
&\sum_{j,k,x,y} \big(\mathsf{M}(j)^\top \otimes \mathsf{M}(k)^\top \otimes |j\rangle\langle j| \otimes |k\rangle\langle k|\big)\,[\Xi^{(2)} \otimes \mathbf{id}_3]\,\big(\mathsf{N}(x)^\top \otimes \mathsf{N}(y)^\top \otimes |x\rangle\langle x| \otimes |y\rangle\langle y|\big) \\
&\quad = \sum_{j,k,x,y} \Big(\big(\mathsf{M}(j)^\top \otimes \mathsf{M}(k)^\top \otimes |j\rangle\langle j|\ \big)\,[\Xi^{(2)}]\,\big(\mathsf{N}(x)^\top \otimes \mathsf{N}(y)^\top \otimes |x\rangle\langle x|\big)\Big) \otimes (|k\rangle\langle k|\,|y\rangle\langle y|) \\
&\quad = \sum_{j,x,y} \Big(\big(\mathsf{M}(j)^\top \otimes \mathsf{M}(y)^\top \otimes |j\rangle\langle j|\ \big)\,[\Xi^{(2)}]\,\big(\mathsf{N}(x)^\top \otimes \mathsf{N}(y)^\top \otimes |x\rangle\langle x|\big)\Big) \otimes |y\rangle\langle y| = O
\end{aligned} \tag{5.7}$$

Since $|y\rangle\langle y|$ are positive operators for both $y = 1$ and $y = 2$, for the above equality to be satisfied, the following equality needs to be satisfied for all y.

$$\sum_{j,x} \big(\mathsf{M}(j)^\top \otimes \mathsf{M}(y)^\top \otimes |j\rangle\langle j|\ \big)\,[\Xi^{(2)}]\,\big(\mathsf{N}(x)^\top \otimes \mathsf{N}(y)^\top \otimes |x\rangle\langle x|\big) = O. \tag{5.8}$$

Now, from our assumption that each of the involved effects is full rank and the consequent existence of the corresponding operator inverses, we can sandwich the above equation, with $\mathbf{id} \otimes (\mathsf{M}(y)^\top)^{-1} \otimes \mathbf{id}$ from left and $\mathbf{id} \otimes (\mathsf{N}(y)^\top)^{-1} \otimes \mathbf{id}$ from right, resulting in

$$\sum_{j,x} \big(\mathsf{M}(j)^\top \otimes \mathbf{id} \otimes |j\rangle\langle j|\ \big)\,[\Xi^{(2)}]\,\big(\mathsf{N}(x)^\top \otimes \mathbf{id} \otimes |x\rangle\langle x|\big) = O. \tag{5.9}$$

Tracing out the system $\mathcal{H}_2$ from the above equation and using the tester normalisation condition $\mathrm{Tr}_2\{\Xi^{(2)}\} = \xi \otimes \mathbf{id}_1$, we have

$$\begin{aligned}
\mathrm{Tr}_2\left\{\sum_{j,x} \big(\mathsf{M}(j)^\top \otimes \mathbf{id} \otimes |j\rangle\langle j|\ \big)\,[\Xi^{(2)}]\,\big(\mathsf{N}(x)^\top \otimes \mathbf{id} \otimes |x\rangle\langle x|\big)\right\} &= O \quad (5.10)\\
\sum_{j,x} \big(\mathsf{M}(j)^\top \otimes |j\rangle\langle j|\ \big)\,[\mathrm{Tr}_2\{\Xi^{(2)}\}]\,\big(\mathsf{N}(x)^\top \otimes |x\rangle\langle x|\big) &= O \quad (5.11)\\
\sum_{j,x} \big(\mathsf{M}(j)^\top \otimes |j\rangle\langle j|\ \big)\,[\xi \otimes \mathbf{id}_1]\,\big(\mathsf{N}(x)^\top \otimes |x\rangle\langle x|\big) &= O \quad (5.12)\\
\sum_{j,x} \big(\mathsf{M}(j)^\top[\xi]\mathsf{N}(x)^\top\big) \otimes (|j\rangle\langle j|\,|x\rangle\langle x|) &= O \quad (5.13)\\
\sum_{x} \big(\mathsf{M}(x)^\top[\xi]\mathsf{N}(x)^\top\big) \otimes |x\rangle\langle x| &= O. \quad (5.14)
\end{aligned}$$

Since $|x\rangle\langle x|$ are positive operators, the last of the above equations warrants that $\big(\mathsf{M}(x)^\top[\xi]\mathsf{N}(x)^\top\big) = O$ for all x. Moreover, following our previous studies, sandwiching these equations with $(\mathsf{M}(x)^\top)^{-1}$ from left and with $(\mathsf{N}(x)^\top)^{-1}$ from right, results in $\xi = O$. This reflects the fact that there does not exist such a binary 2-slot tester that shall achive perfect labelability. Equivalently, if the observable does not admit perfect labelability in a single shot, neither does it in two shots, through adaptive protocols. This completes the proof for two shots of the observable but can be extended to any finite shots. □

Since the above two schemes dealt with all possible scenarios of multiple-shot labeling for binary observables, we can coalesce the above two propositions into the following no-go theorem, whose proof is also a coalition of the above two proofs.

Theorem 5.1.1. If a binary observable $\mathsf{M} \in \mathbf{Obs}(\underline{2}, \mathbb{C}^2)$ does not admit perfect labeling in a single shot, then it does not admit perfect labeling in any finite number of shots.

Achieving perfection through memories

In general, what makes multiple-shot discrimination schemes more powerful is the potential contribution, in performance enhancement, from the quantum memories present in the them. This is also a feature which is absent in schemes which test single-shot scenarios. Regardless, we find that these quantum memories do not help achieve perfect labelability.

5.1.2 Minimum-error labeling

We have found the single-shot optimal minimum-error probability associated with labeling binary observables. Here we will investigate whether this probability is improved when we have access to more number of copies of the unlabeled binary observable and if so, to what degree. We investigate this case when two uses of the observable are available. Here, the labeling experiments are described by 2-slot quantum testers with process effects T_M and T_N and normalisation $\Xi^{(2)} \otimes \mathbf{id}$ with associated causality constraints. Then, the average error probability reads

$$\begin{aligned} p_{\text{error}} =& \frac{1}{2} \left\{ \mathrm{Tr} \left\{ \mathsf{T}_\mathsf{M} \Phi_\mathsf{N}^{\otimes 2} \right\} + \mathrm{Tr} \left\{ \mathsf{T}_\mathsf{N} \Phi_\mathsf{M}^{\otimes 2} \right\} \right\} \\ =& \frac{1}{2} \left\{ \mathrm{Tr} \left\{ \mathsf{T}_\mathsf{M} (\Phi_\mathsf{N}^{\otimes 2} - \Phi_\mathsf{M}^{\otimes 2}) \right\} + \mathrm{Tr} \left\{ (\Xi^{(2)} \otimes \mathbf{id})(\Phi_\mathsf{M}^{\otimes 2}) \right\} \right\} . \end{aligned} \tag{5.15}$$

The Choi-Jamiołkowski operators appearing above can equivalently be represented as,

$$\Phi_{\mathsf{M}}^{\otimes 2} = \sum_i \sum_j \mathsf{M}(i) \otimes |i\rangle\langle i| \otimes \mathsf{M}(j) \otimes |j\rangle\langle j| \,, \tag{5.16}$$

$$\Phi_{\mathsf{N}}^{\otimes 2} = \sum_i \sum_j \mathsf{N}(i) \otimes |i\rangle\langle i| \otimes \mathsf{N}(j) \otimes |j\rangle\langle j| \,. \tag{5.17}$$

Similar to the calculation appearing in the single-shot case, evaluating the term appearing in the error probability, $\mathrm{Tr}\left\{(\Xi^{(2)} \otimes \mathbf{id})(\Phi_{\mathsf{M}}^{\otimes 2})\right\}$, we can find that it is equal to one. This renders the error probability as $p_{\mathrm{error}} = \frac{1}{2}\left\{1 + \mathrm{Tr}\left\{\mathsf{T}_{\mathsf{M}}(\Phi_{\mathsf{N}}^{\otimes 2} - \Phi_{\mathsf{M}}^{\otimes 2})\right\}\right\}$. As such, we have to minimise the term $\mathrm{Tr}\left\{\mathsf{T}_{\mathsf{M}}(\Phi_{\mathsf{N}}^{\otimes 2} - \Phi_{\mathsf{M}}^{\otimes 2})\right\}$. The difference of the operators is given by,

$$\begin{aligned}
(\Phi_{\mathsf{N}}^{\otimes 2} - \Phi_{\mathsf{M}}^{\otimes 2}) &= (M_1 \otimes M_1 - M_2 \otimes M_2) \otimes |1\rangle\langle 1| \otimes |1\rangle\langle 1| + \\
&= (M_1 \otimes M_2 - M_2 \otimes M_1) \otimes |1\rangle\langle 1| \otimes |2\rangle\langle 2| + \\
&= (M_2 \otimes M_1 - M_1 \otimes M_2) \otimes |2\rangle\langle 2| \otimes |1\rangle\langle 1| + \\
&= (M_2 \otimes M_2 - M_1 \otimes M_1) \otimes |2\rangle\langle 2| \otimes |2\rangle\langle 2| \,.
\end{aligned} \tag{5.18}$$

Since M_1 and M_2 commute with each other, the operators $(M_i \otimes M_j - M_k \otimes M_l)$ also commute mutually within themeselves and these operators share the same eigenprojectors $\{|\omega_x\rangle\langle\omega_x|\}$. As such, we can have the following spectral decompositions,

$$\begin{aligned}
(M_1 \otimes M_1 - M_2 \otimes M_2) &= \sum_{x=1}^{d^2} \mu_x \, |\omega_x\rangle\langle\omega_x| \,, \\
(M_1 \otimes M_1 - M_2 \otimes M_2) &= \sum_{x=1}^{d^2} \nu_x \, |\omega_x\rangle\langle\omega_x| \,.
\end{aligned} \tag{5.19}$$

Then, we can rewrite $(\Phi_{\mathsf{N}}^{\otimes 2} - \Phi_{\mathsf{M}}^{\otimes 2})$ as the following,

$$\begin{aligned}
(\Phi_{\mathsf{N}}^{\otimes 2} - \Phi_{\mathsf{M}}^{\otimes 2}) &= (M_1 \otimes M_1 - M_2 \otimes M_2) \otimes \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} + \\
& \quad (M_1 \otimes M_2 - M_2 \otimes M_1) \otimes \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} .
\end{aligned} \tag{5.20}$$

From above, we can read that $(\Phi_{\mathsf{N}}^{\otimes 2} - \Phi_{\mathsf{M}}^{\otimes 2})$ has $4d^2$ real eigenvalues,

$\mu_1, \ldots, \mu_{d^2}, -\mu_1, \ldots, -\mu_{d^2}$ and $\nu_1, \ldots, \nu_{d^2}, -\nu_1, \ldots, -\nu_{d^2}$. We can now write the difference operator as

$$\begin{aligned}(\Phi_{\mathsf{N}}^{\otimes 2} - \Phi_{\mathsf{M}}^{\otimes 2}) \;=\; & \sum_x \mu_x \left|\omega_x \otimes 11\right\rangle\!\left\langle\omega_x \otimes 11\right| - \sum_x \mu_x \left|\omega_x \otimes 22\right\rangle\!\left\langle\omega_x \otimes 22\right| + \\ & \sum_x \nu_x \left|\omega_x \otimes 12\right\rangle\!\left\langle\omega_x \otimes 12\right| - \sum_x \nu_x \left|\omega_x \otimes 21\right\rangle\!\left\langle\omega_x \otimes 21\right| . \end{aligned} \tag{5.21}$$

When we plug-in this operator into the minimum error probability, we get the trace term out to be,

$$\begin{aligned}\mathrm{Tr}\left\{\mathsf{T}_{\mathsf{M}}(\Phi_{\mathsf{N}}^{\otimes 2} - \Phi_{\mathsf{M}}^{\otimes 2})\right\} \;=\; & \sum_x \mu_x \left\langle \omega_x \otimes 11 \middle| \mathsf{T}_{\mathsf{M}} \middle| \omega_x \otimes 11 \right\rangle - \\ & \sum_x \mu_x \left\langle \omega_x \otimes 22 \middle| \mathsf{T}_{\mathsf{M}} \middle| \omega_x \otimes 22 \right\rangle + \\ & \sum_x \nu_x \left\langle \omega_x \otimes 12 \middle| \mathsf{T}_{\mathsf{M}} \middle| \omega_x \otimes 12 \right\rangle - \\ & \sum_x \nu_x \left\langle \omega_x \otimes 21 \middle| \mathsf{T}_{\mathsf{M}} \middle| \omega_x \otimes 21 \right\rangle . \end{aligned} \tag{5.22}$$

For each x, we have four numbers $\mu_x, -\mu_x, \nu_x, -\nu_x$, two of which are positive and the other two negative. We want to sort the each of the summands appearing above into postitive and negative parts. After observing the terms, we can introduce the following "sorting symbols":

$$\begin{aligned} a_x = 1 \otimes 1 \text{ and } b_x = 2 \otimes 2 &: \text{ when } \mu_x \geq 0 \\ a_x = 2 \otimes 2 \text{ and } b_x = 1 \otimes 1 &: \text{ when } \mu_x < 0 \\ c_x = 1 \otimes 2 \text{ and } d_x = 2 \otimes 1 &: \text{ when } \nu_x \geq 0 \\ a_x = 2 \otimes 1 \text{ and } d_x = 1 \otimes 2 &: \text{ when } \nu_x < 0 \end{aligned} \tag{5.23}$$

This enables us to write the trace term as,

$$\begin{aligned}\mathrm{Tr}\left\{\mathsf{T}_{\mathsf{M}}(\Phi_{\mathsf{N}}^{\otimes 2}-\Phi_{\mathsf{M}}^{\otimes 2})\right\} &= \underbrace{\sum_x |\mu_x| \langle \omega_x \otimes a_x|\mathsf{T}_{\mathsf{M}}|\omega_x \otimes a_x\rangle}_{\text{positive}} \\ &\quad \underbrace{-\sum_x |\mu_x| \langle \omega_x \otimes b_x|\mathsf{T}_{\mathsf{M}}|\omega_x \otimes b_x\rangle}_{\text{negative}} \\ &\quad \underbrace{+\sum_x |\nu_x| \langle \omega_x \otimes c_x|\mathsf{T}_{\mathsf{M}}|\omega_x \otimes c_x\rangle}_{\text{positive}} \\ &\quad \underbrace{-\sum_x |\nu_x| \langle \omega_x \otimes d_x|\mathsf{T}_{\mathsf{M}}|\omega_x \otimes d_x\rangle}_{\text{negative}}. \end{aligned} \tag{5.24}$$

Since we want to minimise the value of the trace term, an optimal choice of tester corresponds to suppressing the positive terms as much as possible while preserving the negative ones. This is realised when $\mathsf{T}_{\mathsf{M}}|\omega_x \otimes a_x\rangle = \mathbf{0}$ and $\mathsf{T}_{\mathsf{M}}|\omega_x \otimes c_x\rangle = \mathbf{0}$ are satisfied simultaneously for all x. Consequently, the error probability becomes,

$$\begin{aligned} p_{\text{error}} &= \frac{1}{2} - \frac{1}{2}\sum_x |\mu_x| \langle \omega_x \otimes b_x|\mathsf{T}_{\mathsf{M}}|\omega_x \otimes b_x\rangle - \\ &\qquad \frac{1}{2}\sum_x |\nu_x| \langle \omega_x \otimes d_x|\mathsf{T}_{\mathsf{M}}|\omega_x \otimes d_x\rangle . \end{aligned} \tag{5.25}$$

The minimal value of this expression gives the optimal average error probability and is given by,

$$\begin{aligned} p_{\text{error}} &= \frac{1}{2}(1-\max\{\|M_1\otimes M_1 - M_2\otimes M_2\|_2, \|M_1\otimes M_2 - M_2\otimes M_1\|_2\}) \\ &= \frac{1}{2}(1-\|M_1\otimes M_1 - M_2\otimes M_2\|_2). \end{aligned} \tag{5.26}$$

The first norm is always the maximal value due to the relationship between the eigen-values of the associated effects and the fact that those are probabilities. We can duly repeat this exercise for any finite n copies of the unlabeled observable to have the following result.

> The optimal averaged error with a binary observable, with effects M_1 and M_2, can be labeled with n shots of the observable is given by,
>
> $$p_{\text{error}} = \frac{1}{2}(1-\left\|M_1^{\otimes n} - M_2^{\otimes n}\right\|_2). \tag{5.27}$$

With two shots, we find that it does not improve the single-shot case. As such, we would expect that there will not be any furthur improvements with more number of shots. But, counter-intuitively we find that the error is increasing with more number of shots. We find that this primarily because of the norm appearing, which is the operator 2-norm, as well as the peculiar structure of the labeling problem.

5.1.3 Imperfect unambiguous labeling

We can look into multiple-shot imperfect unambiguous labeling for those binary observables which do not admit imperfect unambiguous labeling with a single shot. These are those observables with non-identical full-rank effects.

Corollary 5.1.1. A binary observable $\mathsf{M} \in \mathbf{Obs}(\underline{2}, \mathcal{H})$, which cannot be imperfect unambiguously labeled in a single-shot, cannot be labeled in multiple-shots.

Proof. As with the above scenarios, let us consider the case when we have two uses of the observable. But, after learning how the case of single-shot went, prior to evaluating the failure probability $p_f = \frac{1}{2}\mathrm{Tr}\left\{\mathsf{T}_?(\Phi_{\mathsf{M}}^{\otimes 2} + \Phi_{\mathsf{N}}^{\otimes 2})\right\}$, let us see whether the no-error conditions can be satisfied by some non-zero T_{M} and T_{N}. For this case, these conditions are

$$\mathrm{Tr}\left\{\Phi_{\mathsf{M}}^{\otimes 2}\mathsf{T}_{\mathsf{N}}\right\} = 0 = \mathrm{Tr}\left\{\Phi_{\mathsf{N}}^{\otimes 2}\mathsf{T}_{\mathsf{M}}\right\}. \tag{5.28}$$

We can note that,

$$\tilde{\Phi}_{\mathsf{M}}^{\otimes 2} = \begin{pmatrix} M_1 & O & O & O \\ O & M_2 & O & O \\ O & O & M_1 & O \\ O & O & O & M_2 \end{pmatrix}. \tag{5.29}$$

Since M_1 and M_2 are full-rank operators, $\Phi_{\mathsf{M}}^{\otimes 2}$ is also a full-rank operator. As such, the only condition for one of the no-error conditions to be satisfied is to satisfy $\mathsf{T}_{\mathsf{N}} = O$. This implies that there does not exist a non-zero T_{N} and consequently a non-trivial tester which could carry out this labeling experiment. □

No perfect, then nothing at all!

Through the due course of this thesis up until now, we have learnt that a binary observable that does not admit perfect labeling in a single-shot neither will admit imperfect unambiguous labeling. Subsequently, even when we came to the multiple-shot regime, we found that it will neither allow perfect nor imperfect unambiguous labeling. As such, we can conclude that if a binary observable is not perfectly labelable in a single-shot, then the next best labeling experiment we can carry out is minimum-error.

5.2 Non-binary observables

We have learnt that the results for binary observables in the multiple-shot regime are "negative" in the sense that not only we do not witness any performance improvement with more number of shots but see it diminishing. Even with this finding, we would still expect that multiple-shot shall prove to be useful for non-binary observables. This intuition is justified right away as we start to investigate the case for perfect labeling.

5.2.1 Perfect labeling

When it comes to perfect labeling of non-binary observables, we are more interested in the question on what is the minimal number of uses or queries of the unlabeled observable required to perfectly label it. Addressing this question leads us to the following theorem.

Theorem 5.2.1. A non-binary observable can be perfectly labled, using the simple scheme, in $(n-1)$ shots if and only if there exists at least $(n-1)$ effects possessing at least one eigen-value 1.

Proof. To perfectly label a specific effect M_x corresponds to the existence of a quantum state $|\varphi_x\rangle$ such that $M_y|\varphi_x\rangle = \mathbf{0}$ for all $x \neq y$ and $M_x|\varphi_x\rangle \neq \mathbf{0}$. This translates to $\sum_{x\neq y} M_y = \mathbf{0}$. Using the normalisation condition for the observable, $\sum_y M_y = \mathbf{id}$, we have $\sum_y M_y|\varphi_x\rangle = |\varphi_x\rangle$. This is nothing but $M_x|\varphi_x\rangle + \sum_{y\neq x} M_y|\varphi_x\rangle = |\varphi_x\rangle$, implying that $M_x|\varphi_x\rangle = |\varphi_x\rangle$. This imples, in turn, that $|\varphi_x\rangle$ needs to be an eigen-vector of M_x with associated eigen-value 1. Now, for a non-binary observable to be perfectly labled, based on this simple scheme, at least $(n-1)$ effects need to be perfectly labeled. Consequently, at least $(n-1)$ effects should have at least one eigen-value 1. $\square$

We note that such a simple scheme is possible if there exists an appropriate collection of probe states $(|\varphi_1\rangle, \ldots, |\varphi_{(n-1)}\rangle)$ and we measure each shot of the unlabeled observable on these states.

On binarisation

Binarisation of an non-binary observable is a procedure in which we reduce the number of outcome labels to two. Essentially, we are constructing a binary observable out of a non-binary one. The binarsiation we consider here is the following specific one: Consider a non-binary observable $\mathsf{M} \in \mathbf{Obs}(\underline{n}, \mathcal{H})$, with effects

$\mathsf{M}(1) = M_1, \mathsf{M}(2) = M_2, \dots \mathsf{M}(n) = M_n$. We can "binarise" this observable to another $\mathsf{N} \in \mathbf{Obs}(\underline{2}, \mathcal{H})$ by the consideration $\mathsf{N}(1) = M_1$ and $\mathsf{N}(2) = M_2 + \cdots + M_n$. We can also note that given such a non-binary observable, we can have $\binom{n}{1} = n$ binarisations. Now, equipped with this notion, we can have the following proposition.

Proposition 5.2.1. If a non-binary observable can be perfectly labeled using the simple scheme, then each binarisation (composed of two effects of the form, M_x and $\sum_{y \neq x} M_y$) can be perfectly labeled in a single-shot.

Proof. We have seen that a binary observable can be perfectly labeled if and only if at least one of the effects is rank-deficient. And, we have just observed that there are n possible binarisations (of the above specific construction). If we assume that our observable can be labeled perfectly in the simple scheme as discussed above, it should have at least $(n-1)$ with at least one eigen-value 1. Consequently, in any such binarisation, at least one of the effects is rank deficient, thus entailing single shot perfect labelability to each of them. $\square$

Form of effects

The effects of those observables $\mathsf{M} \in \mathbf{Obs}(\underline{n}, \mathcal{H})$ with $\mathbf{dim}(\mathcal{H}) \geq n$, have the following form,

$$M_x = |b_x\rangle\langle b_x| + A_x \quad \text{for all } x \in [1, (n-1)] \text{ and } M_n = A_n, \tag{5.30}$$

where $\{|b_x\rangle\}_{x=1}^{n-1}$ is a orthonormal basis for an $(n-1)$ dimensional subspace $\mathcal{H}_{(n-1)}$ are A_x are positive semi-definite operators satisfying $\sum_{x=1}^{n} A_x = \mathbf{id}_{\mathcal{H} \setminus \mathcal{H}_{(n-1)}}$.

Observables with identical effects

When observables have identical collections of effects composing, the above theorem does not apply strictly. For example, consider the observable composed of effects $M_1 = |\varphi\rangle\langle\varphi|$, $M_2 = \cdots = M_n = \frac{1}{(n-1)}(\mathbf{id} - |\varphi\rangle\langle\varphi|)$. It is sufficient to use a single use of the observable to perfectly label it; Measure the observable on the state $|\varphi\rangle$ and the recorded outcome label is associated with the effect $|\varphi\rangle$ and rest of the effects all correspond to $(\mathbf{id} - |\varphi\rangle\langle\varphi|)$. Such pathological examples can be designed with ease where the observables in question can be perfectly labeled in $m < n$ uses.

5.3 Partial labeling and Antilabeling

We continue the discussion on partial and antilabeling to cases within multiple-shot regime.

5.3.1 Partial labeling

Given a non-binary observable $\mathsf{M} \in \mathbf{Obs}(\underline{n}, \mathcal{H})$, with effects $M_1, \cdots M_n$, let us assume that there exist two non-identical strict subsets of effects X_{ϕ_1} and X_{ϕ_2} such that $M_{\phi_1} |\phi_1\rangle = |\phi_1\rangle$ and $M_{\phi_2} |\phi_2\rangle = |\phi_2\rangle$ for two non-identical states $|\phi_1\rangle$ and $|\phi_2\rangle$, where $M_{\phi_1} = \sum_{j \in X_{\phi_1}} M_j$. Then we can have a two-shot partial labeling scheme as follows: Measure the unlabeled observable on the state $|\phi_1\rangle$. The recorded outcome label is associated with an effect $M_x \in X_{\phi_1}$. In the second shot, measure the observable on the state $|\phi_2\rangle$ and the recorded label is associated to $M_y \in X_{\phi_2}$. Note that we have not assumed any restrictions on the disjointness between sets X_{ϕ_1} and X_{ϕ_2}. Assume that they are not disjoint. Then, it could happen that in both shots, the same label is recorded. In a case, we can conclude that the associated effect belongs to X_{ϕ_1} and X_{ϕ_2}, thus reducing the possibilities furthur. Moreove, when a set X_ϕ is singleton, you identify the recorded label perfectly. This scheme of partial labeling can be presented as the following.

> Given a non-binary rank-deficient observable $\mathsf{M} \in \mathbf{Obs}(\underline{n}, \mathcal{H})$, composed of effects $M_1, \ldots, M_n$, there exists m non-identical strict subsets X_{ϕ_j} such that $M_{\phi_j} |\phi_j\rangle = |\phi_j\rangle$. Then, we can have a m-shot scheme that partially labels the observable: Measure the observable on $|\phi_1\rangle, \ldots, |\phi_m\rangle$ and the respective recorded outcome labels are associated with effects $M_x \in X_{\phi_1}, \ldots, M_y \in X_{\phi_m}$. If a same label is recorded for states $|\phi_a\rangle, |\phi_b\rangle, \ldots,$ then the label is associated with an effect belonging to the intersection of the associated sets $\cap_{\alpha=a,b,\ldots} X_{\phi_\alpha}$.

5.3.2 Antilabeling

With partial lableling discussed as it is above, we note that antilabeling is "complementary" to partial labeling in the following sense. Here, we define $\tilde{M}_{\phi_x} = \sum_{j \in X_{\phi_x}} M_j$ such that $\tilde{M}_{\phi_x} |\phi_x\rangle = \mathbf{0}$. Let us see a two-shot case: There exists at least two sets X_{ϕ_1} and X_{ϕ_2} such that $\tilde{M}_{\phi_1} |\phi_1\rangle = \mathbf{0}$ and $\tilde{M}_{\phi_2} |\phi_2\rangle = \mathbf{0}$. Measuring the observable on the state $|\phi_1\rangle$, we can conclude that the recorded label does not associate to any effect from X_{ϕ_1}. With a subsequent measurement on the state $|\phi_2\rangle$, we can conclude that the label is not associated with any effect from

X_{ϕ_2}. If it happens that the same label is recorded for both of the shots, then we can assert that the label does not associate to any effect from $X_{\phi_1} \cup X_{\phi_2}$. This scheme can be written down as the following.

> Given a non-binary rank-deficient observable $\mathsf{M} \in \mathbf{Obs}(\underline{n}, \mathcal{H})$, composed of effects $M_1, \ldots, M_n$, there exists m non-identical strict subsets X_{ϕ_j} such that $\tilde{M}_{\phi_j} |\phi_j\rangle = |\phi_j\rangle$. Then, we can have a m-shot scheme that antilabels the observable: Measure the observable on $|\phi_1\rangle, \ldots, |\phi_m\rangle$ and the respective recorded outcome labels are not associated with effects $M_x \in X_{\phi_1}, \ldots, M_y \in X_{\phi_m}$. If a same label is recorded for states $|\phi_a\rangle, |\phi_b\rangle, \ldots,$ then the label is not associated with an effect belonging to the intersection of the associated sets $\cup_{\alpha=a,b,\ldots} X_{\phi_\alpha}$.

Chapter 6

Advantage of higher-order incompatibility

As discussed in the introductory chapter, incompatibility of observables has been identified as a genuine quantum resource in an array of quantum information processing tasks. In this chapter, we investigate the resourcefulness of incompatible testers in distinguishability tasks concerning quantum combs. To aid this investigation, we introduce a two-party comb distinguishability game, which is a generalised version to the one presented in [SŠC19]. Now, prior to sketching this game and analysising it, we can begin this chapter by discussing relevant notions regarding incompatible quantum testers.

6.1 Higher-order incompatibility

In this section we discuss the notion of compatibility, or equivalently incompatibility, of quantum testers and associated characterisations. Within the scope of this thesis, whereever we refer to "higher-order incompatiblity" we refer to incompatibility of testers. In fact, there could be other higher-order quantum theoretic notions of incompability which could include joint implementation of quantum combs for example; But at this point of writing, there are not any notion other than tester incompatibility that has been introduced and developed. Tester incompatibility was introduced and charactersied by Sedlák *et.al.* in [SRCZ16] which serves as the prime source for this discussion.

6.1.1 Incompatible testers

Compatibility of quantum testers is defined in a similar way to that of quantum observables. Let $\mathsf{S} \in \mathbf{Tester}(\underline{m}, N)$ and $\mathsf{T} \in \mathbf{Tester}(\underline{n}, N)$. Suppose the following are the statistics produced by these two testers, respectively, when an appropriate comb Φ is measured by them.

$$p_j = \mathrm{Tr}\left\{\mathsf{S}_j \Phi\right\} \quad \text{and} \quad q_k = \mathrm{Tr}\left\{\mathsf{T}_k \Phi\right\}. \tag{6.1}$$

These two testers S and T are compatible if there exists at least one tester $\mathsf{R} \in \mathbf{Tester}(\underline{o}, N)$ such that the above statistics can be reproduced separately by postprocessing the statistics R produces on any arbitrary Φ. This translates to existence of postprocessing matrices ν_{xy} and μ_{ab} with the following test effects satisfying,

$$\mathsf{S}_x = \sum_{y \in \underline{o}} \nu_{xy} \mathsf{R}_y \quad \text{and} \quad \mathsf{T}_a = \sum_{b \in \underline{o}} \mu_{ab} \mathsf{R}_b. \tag{6.2}$$

If there do not exist such a "parent" tester R and associated postprocessings, then S and T are said to be incompatible. This notion of compatibility can be extended to any finite collection of testers, having identical sequences of associated systems. As such, from now onwards, whenever we refer to collections of testers, they are of appropriate mixable ones.

Characterisation

Let $\{\mathsf{S}_k\}$ be a finite collection of testers with respective normalisations $\{\Theta_k\}$ and associated canonical observables $\{\mathsf{M}_k\}$. $\{\mathsf{S}_k\}$ is compatible if and only if the following are satisfied,

- All of the normalisations coincide with each other, $\Theta_i = \Theta_k$, for all k and i.
- The collection of canonical observables, $\{\mathsf{M}_k\}$, is compatible.

6.1.2 Robustness of incompatibility

A characterisation of the "amount" of incompatibility contained in a given incompatbile collection of testers $\{\mathsf{T}_k\}$ is obtained by gauging the minimal noise that is to be added so as to break the incompatibility or "compatibilise" the collection. This noise additions are realised by physically mixing another collection of "noise testers", $\{\mathsf{N}_k\}$, to the incompatible ones. This characterisation is referred to as "robustness of incompatibility" of the given collection $\mathcal{R}(\{\mathsf{T}_k\})$ and is given by,

$$\begin{aligned} \mathcal{R}(\{\mathsf{T}_k\}) &= \text{minimise } r, \\ &\quad \text{such that } \left\{\frac{1}{(1+r)}(\mathsf{T}_k + r\mathsf{N}_k)\right\} \text{ is compatible.} \end{aligned} \tag{6.3}$$

This minimisation is realised by searching over all possible collections of non-identical noise testers. By construction $r = 0$ when we have a compatible collection of testers.

6.2 Quantum comb distinguishability game

We sketch our comb distinguishability game as the following. Let us denote an ensemble of m N-slot quantum combs with $\mathcal{X} = \{(\Phi_b, w(b))\}_{b=1}^{m}$, where, as we have established in the previous chapters, $w(b)$ is the probability with with the comb Φ_b is chosen from the ensemble. Calling the two parties involved Alice and Bob, we can now sketch out a single-round out of the the multi-round distinguishability game.

- Bob has access to a finite collection of s non-identical N-slot comb ensembles, $\mathcal{X}_1, \ldots, \mathcal{X}_s$, with $\mathcal{X}_\beta = \{(\Phi_{b|\beta}, w(b|\beta))\}_{b=1}^{m}$ where $\beta = 1, \ldots, s$. As an illustration, $\Phi_{2|5}$ is the second comb from the fifth ensemble and $w(2|3)$ is the probability with which the second comb is chosen provided the fifth ensemble is already chosen, that is, the associated probability with which $\Phi_{2|5}$ is selected.

- In each round of the protocol, Bob chooses one among the s ensembles with probability $w(\beta)$. From the chosen ensemble $\mathcal{X}_\beta$, one among the m combs $\Phi_{b|\beta}$ is chosen with the conditional probability $w(b|\beta)$.

- Bob sends the classical information β, regarding which ensemble he chose as well as a black box that implements the comb $\Phi_{b|\beta}$ to Alice.

- Alice is now tasked to guess which comb the received black box will implement, provided that she has been reported the value of β. Essentially, she has to guess the value of b.

- Alice has access to fixed, finite collection of appropriate $(N+1)$-slot testers $\{\mathsf{T}_\alpha\}_\alpha$. She chooses and implements a tester from this collection on the comb she has been given. This measurement process results in an outcome a which occurs with the conditional probability $p(a|b) = \mathrm{Tr}\left\{\Phi_{b|\beta}\mathsf{T}_{a|\alpha}\right\}$.

Now, we want to compare between optimal performances furnished by incompatible and compatible testers. That is, to compare the two possible scenarios of the game; One where she plays it while she has restricted access to incompatible collections of testers and the other where she does this with compatible ones.

Gameplay with incompatible testers

In this scenario of gameplay, the collection of testers $\{\mathsf{T}_\alpha\}_\alpha$ available to Alice is restricted to be an incompatible collection. She also has access to a classical random variable λ. In each round, after receiving Bob's choice of the ensemble β and the comb black box, Alice chooses and implements the tester T_α, out of the collection $\{\mathsf{T}_\alpha\}_\alpha$, with the conditional probability $p(\alpha|\beta,\lambda)$ as prescribed by the random variable λ. Implementation of the tester results in a measurement outcome a. After acquiring the outcome a, Alice makes an educated guess of which comb it was, that is, she guesses the value of b as g. We notice that in this gameplay, Bob's ensembles and Alice's incompatible collection of testers are fixed for the game. As such, the averaged success probability of guessing is given by

$$p_s(\{\mathcal{X}_\beta\},\{\mathsf{T}_\alpha\},S) = \tag{6.4}$$

$$= \sum_{b,\beta,a,\alpha,g,\lambda} w(\beta)w(b|\beta)p(\lambda)p(\alpha|\beta,\lambda)\times \mathrm{Tr}\left\{\Phi_{b|\beta}\mathsf{T}_{a|\alpha}\right\} p(g|a,\beta,\lambda)\delta_{g,b} \tag{6.5}$$

$$= \sum_{b,\beta,a,\alpha,g,\lambda} w(b,\beta)p(\lambda)p(\alpha|\beta,\lambda)\times \mathrm{Tr}\left\{\Phi_{b|\beta}\mathsf{T}_{a|\alpha}\right\} p(g|a,\beta,\lambda)\delta_{g,b}. \tag{6.6}$$

Here, we have written the product probability $w(\beta)w(b|\beta)$ as the joint probability $w(b,\beta)$. The Kronecker delta functions $\delta_{g,b}$ ensure that we are taking into account only events of successful guesses. Here, the variables she has control over are the random variable λ, the model based on which she chooses the tester and, the model with which she makes guesses. This triple of variables prescribes a strategy and, from a point of view of optimising the guessing probability, such strategies can be expressed as to constitite the set of strategies, $\mathcal{S} = \{(p(\lambda), p(\alpha|\beta,\lambda), p(g|a,\beta,\lambda))\}$. As such, the third argument appearing in p_s is $S\in\mathcal{S}$. In order to maximise the above success probability, Alice needs to optimise over all possible strategies she has access to. Thus, the optimal success probability is given by optimising over all such strategies,

$$p_s(\{\mathcal{X}_\beta\},\{\mathsf{T}_\alpha\}) = \tag{6.7}$$

$$\max_{\mathcal{S}} \sum_{b,\beta,a,\alpha,g,\lambda} w(b,\beta)p(\lambda)p(\alpha|\beta,\lambda)\times \mathrm{Tr}\left\{\Phi_{b|\beta}\mathsf{T}_{a|\alpha}\right\} p(g|a,\beta,\lambda)\delta_{g,b}. \tag{6.8}$$

Gameplay with compatible testers

In this scenario of gameplay, Alice has restricted access to only a fixed tester Q_ν that she can implement per a whole run of the game. Despite this, on different runs of the game (not to be confused with different rounds of a single run of the game) she can choose different testers with the help of the classical random variable ν. That is, she chooses the fixed tester Q_ν with probability $p(\nu)$. Moreover, having access to a single fixed tester Q_ν is equivalent to having access to a finite compatible collection of testers, whose joint tester is none other than Q_ν. Now, the average success probability is given by

$$p_s(\{\mathcal{X}_\beta\}, \mathsf{Q}_\nu, S_c) = \sum_{b,\beta,a,g,\nu} w(b,\beta)p(\nu) \times \mathrm{Tr}\left\{\Phi_{b|\beta}\mathsf{Q}_{a|\nu}\right\} p(g|a,\beta,\nu)\delta_{g,b}. \tag{6.9}$$

Here, we can notice that the tester Q_ν can be any possible tester, each run of the game. This is in contrast to the previous gameplay where the available collection of testers is fixed for all runs of the game. So, Alice can cleverly choose one each run of the game. As such, the choice of tester as well as the model used to make the choice become part of Alice's strategy. Moreover, the model she uses to make her guesses is also essentially part of her strategy. Thus, we have the set of triples $\mathcal{S}_C = \{(p(\nu), \mathsf{Q}_\nu, p(g|a,\beta,\lambda))\}$ as the strategies available to Alice; In the above equation $S_c \in \mathcal{S}_C$. Thus, the optimal success probability is given by optimising over all such strategies,

$$p_s(\{\mathcal{X}_\beta\}) = \max_{\mathcal{S}_C} \sum_{b,\beta,a,g,\nu} w(b,\beta)p(\nu) \times \mathrm{Tr}\left\{\Phi_{b|\beta}\mathsf{Q}_{a|\nu}\right\} p(g|a,\beta,\nu)\delta_{g,b}. \tag{6.10}$$

The advantage

We have seen that how Alice can play the game through two contrasting scenarios. To gauge the performance increase achieved, by the gameplay with incompatible testers over that by the one with compatible testers, we can use the relative quantity,

$$\frac{p_s(\{\mathcal{X}_\beta\}, \{\mathsf{T}_\alpha\})}{p_s(\{\mathcal{X}_\beta\})}. \tag{6.11}$$

Theorem 6.2.1. When optimised over all possible games (Bob's collection of ensembles), the associated maximal relative optimal success probability, with any particular collection of incompatible testers $\{\mathsf{T}_\alpha\}$ over all possible compatible collections, is given by

$$\max_{\{\mathcal{X}_\beta\}} \frac{p(\{\mathcal{X}_\beta\}, \{\mathsf{T}_\alpha\})}{p(\{\mathcal{X}_\beta\})} = 1 + \mathcal{R}(\{\mathsf{T}_\alpha\}). \tag{6.12}$$

The proof of this theorem is the discussion of the subsequent section.

6.3 SDP

In this section, we write down the SDP for evaluating the robustness of incompatibility (ROI) for a collection of testers $\{\mathsf{T}_k\}$. Prior to that, we look at the ROI tailored for our game.

Robustness of incompatibility

The ROI is evaluated to see when the convex mixture $\frac{1}{(1+r)}(\mathsf{T}_{a|\alpha} + r\mathsf{N}_{a|\alpha})$ is compatible. This corresponds to existence of noise testers $\{\mathsf{N}_\alpha\}_\alpha$, classical postprocessing $p(a|\alpha,\mu)$ and parent observable Q_μ. ROI is essentially rooting out the minimal r over such "compatibilising" triples. We write the ROI here and the sketch out all the objects involved in detail.

$$\begin{aligned} \mathcal{R}(\{\mathsf{T}_\alpha\}) \;=\;& \min \; r \\ & \text{s.t. } \frac{1}{(1+r)}(\mathsf{T}_{a|\alpha} + r\mathsf{N}_{a|\alpha}) = \sum_\mu p(a|\alpha,\mu)\mathsf{Q}_\mu \end{aligned} \tag{6.13}$$

$$\begin{aligned}
\text{[principal testers]} \quad & \mathsf{T}_{a|\alpha} \geq 0, \sum_\alpha \mathsf{T}_{a|\alpha} = \Theta_\alpha^{(n)} \otimes \mathbf{id}_{(2n-1)} \\
& \mathrm{Tr}_{(2k-2)}\{\Theta_\alpha^{(n)}\} = \Theta_\alpha^{(k-1)} \otimes \mathbf{id}_{(2k-3)}; k \in \{2,\dots,n\} \\
& \mathrm{Tr}\{\Theta_\alpha^{(1)}\} = 1. \\[1em]
\text{[noise testers]} \quad & \mathsf{N}_{a|\alpha} \geq 0, \sum_\alpha \mathsf{T}_{a|\alpha} = \Gamma_\alpha^{(n)} \otimes \mathbf{id}_{(2n-1)} \\
& \mathrm{Tr}_{(2k-2)}\{\Gamma_\alpha^{(n)}\} = \Gamma^{(k-1)} \otimes \mathbf{id}_{(2k-3)}; k \in \{2,\dots,n\} \\
& \mathrm{Tr}\{\Gamma_\alpha^{(1)}\} = 1. \\[1em]
\text{[post-processing]} \quad & p(a|\alpha,\mu) \geq 0, \sum_\mu p(a|\alpha,\mu) = 1 \\[1em]
\text{[parent tester]} \quad & \mathsf{Q}_\mu \geq 0, \sum_\mu \mathsf{Q}_\mu = \Theta^{(n)} \otimes \mathbf{id} \\
& \mathrm{Tr}_{(2k-2)}\{\Theta^{(n)}\} = \Theta^{(n-1)} \otimes \mathbf{id}_{(2k-3)}; k \in \{2,\dots,n\} \\
& \mathrm{Tr}\{\Theta^{(1)}\} = 1. \\[1em]
\text{[normalisations]} \quad & \Theta^{(n)} = \frac{1}{(1+r)}(\Theta_\alpha^{(n)} + r\Gamma_\alpha^{(n)}) \quad \forall\alpha\,\forall n \\
& \Theta_\alpha^{(n)} \geq 0 \;\text{ and }\; \Gamma_\alpha^{(n)} \geq 0 \quad \forall\alpha\,\forall n.
\end{aligned}$$

6.3.1 Primal SDP

With respect to Eq.(6.14), let r_0 be the robustness of the given collection of testers $\{\mathsf{T}_\alpha\}$, that is, $\mathcal{R}(\{\mathsf{T}_\alpha\}) = r_0$. The process effects of the involved noise testers can be retrieved through

$$\mathsf{N}_{a|\alpha} = \frac{1}{r_0}\{(1+r_0)\sum_\mu p(a|\alpha,\mu)\mathsf{Q}_\mu - \mathsf{T}_{a|\alpha}\}. \tag{6.14}$$

Since $\mathsf{N}_{a|\alpha}$ are positive semidefinite operators, their positive semidefiniteness, $\mathsf{N}_{a|\alpha} \geq O$, translate as follows.

$$\frac{1}{r_m}\{(1+r_0)\sum_\mu p(a|\alpha,\mu)\mathsf{Q}_\mu - \mathsf{T}_{a|\alpha}\} \;\; \geq \;\; O \tag{6.15}$$

$$\implies \underbrace{s}_{:=(1+r_0)} \sum_\mu p(a|\alpha,\mu)\mathsf{Q}_\mu \;\; \geq \;\; \mathsf{T}_{a|\alpha}. \tag{6.16}$$

Assume a_k to be the outcome for the k^{th} tester. Since there are m testers, we can collectively write down the outcomes of each tester as vector string $\vec{a} = (a_1, \ldots, a_m)$. Each of the m testers has o outcomes, that is, $a_k \in \underline{o}$ for all k. Consequentially, there are o^m possible $\vec{a}$ vectors. Now, given the outcome μ of the joint tester, the probability of producing a specific $\vec{a}$ is given by,

$$p(\vec{a}|\mu) = p(a_1|\alpha=1,\mu)\cdot p(a_2|\alpha=2,\mu)\cdots p(a_m|\alpha=m,\mu) = \prod_\alpha p(a_\alpha|\alpha,\mu). \tag{6.17}$$

Given the outcome of the parent tester is μ, the probability of the α^{th} tester producing an outcome equal to a, that is, $a_\alpha = a$ is given by $p(a|\alpha,\mu) = \sum_{\vec{a}} d_{\vec{a}}(a|\alpha)p(\vec{a}|\mu)$. Then,

$$\sum_\mu p(a|\alpha,\mu)\mathsf{Q}_\mu = \sum_{\vec{a}} d_{\vec{a}}(a|\alpha)\mathsf{Q}_{\vec{a}}. \tag{6.18}$$

Primal SDP

$$\Theta^{(k)} = \frac{1}{s}(\Theta^{(k)}_\alpha + r\Gamma^{(k)}_\alpha) \quad \forall\alpha\,\forall k. \tag{6.19}$$

In fact, the above conditions are automatically satisfied since post-processings preserve normalisations and as such does not appear in the SDP. Now, we can formulate the SDP in the following form.

$$\begin{aligned}
1+\mathcal{R}(\{\mathsf{T}_\alpha\}) \;\; &= \;\; \min \;\; s && (6.20)\\
&\text{s.t.} \;\; \sum_{\vec{a}} d_{\vec{a}}(a|\alpha)\tilde{\mathsf{Q}}_{\vec{a}} \geq \mathsf{T}_{a|\alpha} && (6.21)\\
&\tilde{\mathsf{Q}}_{\vec{a}} \geq 0, \; \sum_{\vec{a}} \tilde{\mathsf{Q}}_{\vec{a}} = \mathbf{id}\otimes\tilde{\Theta}^{(n)} && (6.22)\\
&\mathrm{Tr}_{(2k-2)}\{\tilde{\Theta}^{(k)}\} = \mathbf{id}_{(2k-3)}\otimes\tilde{\Theta}^{(k-1)}; \;\; \forall \quad k\in\{2,\ldots,n\} && (6.23)
\end{aligned}$$

Upper bound

In this discussion, we can the primal SDP to upper bound $\frac{p_s(\{\mathcal{X}_\beta\},\{\mathsf{T}_\alpha\})}{p_s(\{\mathcal{X}_\beta\})}$. In order to find this bound, let us first assume that the triple $(N^*_{a|\alpha}, Q^*_\mu, p^*(a|\alpha,\mu))$ is one of the solutions to the primal SDP. This essentially means that the following expressions hold.

$$\frac{1}{1+\mathcal{R}(\{\mathsf{T}_\alpha\})}\left(\mathsf{T}_{a|\alpha}+\mathcal{R}(\{\mathsf{T}_\alpha\})N^*_{a|\alpha}\right)=\sum_\mu p^*(a|\alpha,\mu)Q^*_\mu, \tag{6.24}$$

$$\left(\mathsf{T}_{a|\alpha}+\mathcal{R}(\{\mathsf{T}_\alpha\})N^*_{a|\alpha}\right)=\left(\sum_\mu p^*(a|\alpha,\mu)Q^*_\mu\right)(1+\mathcal{R}(\{\mathsf{T}_\alpha\})). \tag{6.25}$$

Since $\mathcal{R}(\{\mathsf{T}_\alpha\})$ is a positive number and $N^*_{a|\alpha}$ are positive semidefinite operators, $\mathcal{R}(\{\mathsf{T}_\alpha\})N^*_{a|\alpha}$ are also positive semidefinite operators. As such, we have the following operator inequalities for all a and α.

$$(1+\mathcal{R}(\{\mathsf{T}_\alpha\}))\sum_\mu p^*(a|\alpha,\mu)Q^*_\mu \geq \mathsf{T}_{a|\alpha}. \tag{6.26}$$

Now, let us multiply this inequality with $\Phi_{b|\beta}$, the probabilities $w(b,\beta), p(\nu), p(\alpha|\beta,\nu), p(g|a,\beta,\nu)$ and $\delta_{g,b}$. Moreover, by adding over all the appearing indices and taking trace of the resultant operator inequality, we arrive at the inequality,

$$\begin{aligned}(1+\mathcal{R}(\{\mathsf{T}_\alpha\}))&\sum_{\mu,g,a,b,\alpha,\beta,\nu} w(b,\beta)p(\nu)p(\alpha|\beta,\nu)p^*(a|\alpha,\mu)\mathrm{Tr}\left\{\Phi_{b|\beta}Q^*_\mu\right\}p(g|a,\beta,\nu)\delta_{g,b}\\ &\geq \sum_{\mu,g,a,b,\alpha,\beta,\nu} w(b,\beta)p(\nu)p(\alpha|\beta,\nu)\mathrm{Tr}\left\{\Phi_{b|\beta}\mathsf{T}_{a|\alpha}\right\}p(g|a,\beta,\nu)\delta_{g,b}.\end{aligned} \tag{6.27}$$

If we were to have the definition $p(g|\mu,\beta,\nu)=\sum_{a,\alpha}p^*(a|\alpha,\mu)p(\alpha|\beta,\nu)p(g|a,\beta,\nu)$, the above inequality can be written as,

$$\begin{aligned}(1+\mathcal{R}(\{\mathsf{T}_\alpha\}))&\sum_{\mu,g,a,b,\alpha,\beta,\nu} w(b,\beta)p(\nu)\mathrm{Tr}\left\{\Phi_{b|\beta}Q^*_\mu\right\}p(g|\mu,\beta,\nu)\delta_{g,b}\\ \geq&\sum_{\mu,g,a,b,\alpha,\beta,\nu} w(b,\beta)p(\nu)p(\alpha|\beta,\nu)\mathrm{Tr}\left\{\Phi_{b|\beta}\mathsf{T}_{a|\alpha}\right\}p(g|a,\beta,\nu)\delta_{g,b}.\end{aligned} \tag{6.28}$$

The sum, $\sum_{\mu,g,a,b,\alpha,\beta,\nu}w(b,\beta)p(\nu)\mathrm{Tr}\left\{\Phi_{b|\beta}Q^*_\mu\right\}p(g|\mu,\beta,\nu)\delta_{g,b}$, is similar to the the average success probabilities, $\sum_{b,\beta,a,g,\nu}w(b,\beta)p(\nu)\times\mathrm{Tr}\left\{\Phi_{b|\beta}\mathsf{Q}_{a|\nu}\right\}p(g|a,\beta,\nu)\delta_{g,b}$, characterising the gameplay with compatible testers. One notable difference is that unlike $\mathsf{Q}_{a|\nu}$, Q^*_μ does not depend upon the model ν and consequently $p(\nu)$. We can maximise both of these sums over strategies accessible in the compatible gameplay S_C. Following our previous discussion, the corresponding maximum of the later sum is $p_s(\{\mathcal{X}_\beta\})$. Since the earlier sum is more restrictive, essentially due to its independence of ν, it is never larger than $p_s(\{\mathcal{X}_\beta\})$. Thus, we can replace the earlier sum with $p_s(\{\mathcal{X}_\beta\})$ in the above inequality and preserve the direction of the inequality,

$$\begin{aligned}(1+\mathcal{R}(\{\mathsf{T}_\alpha\}))\, p_s(\{\mathcal{X}_\beta\}) \\ \geq \sum_{\mu,g,a,b,\alpha,\beta,\nu} w(b,\beta)p(\nu)p(\alpha|\beta,\nu)\mathrm{Tr}\left\{\Phi_{b|\beta}\mathsf{T}_{a|\alpha}\right\}p(g|a,\beta,\nu)\delta_{g,b}.\end{aligned} \tag{6.29}$$

This inequality is valid for all probability distributions, $p(\nu), p(\alpha,\beta,\nu)$ and $p(g|a,\beta,\nu)$. Consequently, its validity is preserved while me maximise over all such distribution. This maximisation results the right hand side of the inequality to be equal to the optimal success probability achieved in the game with with incompatible gameplay, $p(\{\mathcal{X}_\beta\},\{\mathsf{T}_\alpha\})$. Then maximising over all games also preserve the inequality direction, resulting in

$$1+\mathcal{R}(\{\mathsf{T}_\alpha\}) \geq \max_{\{\mathcal{X}_\beta\}} \frac{p(\{\mathcal{X}_\beta\},\{\mathsf{T}_\alpha\})}{p(\{\mathcal{X}_\beta\})}. \tag{6.30}$$

6.3.2 Dual SDP

Since we have the primal SDP, we can write down the Lagrangian $\mathcal{L}$.

Lagrangian

The Lagrangian associated with our primal SDP is the following.

$$\begin{aligned}\mathcal{L} = \quad & s + \textstyle\sum_{a,\alpha}\mathrm{Tr}\left\{\omega_{a\alpha}\left(\mathsf{T}_{a|\alpha} - \sum_{\vec{a}} d_{\vec{a}}(a|\alpha)\tilde{\mathsf{Q}}_{\vec{a}}\right)\right\} - \mathrm{Tr}\left\{X\left(\mathbf{id}\otimes\tilde{\Theta}^{(n)} - \sum_{\vec{a}}\tilde{\mathsf{Q}}_{\vec{a}}\right)\right\} \\ & -\mathrm{Tr}\left\{\textstyle\sum_{\vec{a}} x_{\vec{a}}\tilde{\mathsf{Q}}_{\vec{a}}\right\} - \mathrm{Tr}\left\{y^{(k)}(\mathbf{id}_{(2k-3)}\otimes\tilde{\Theta}^{(k-1)} - \mathrm{Tr}_{(2k-2)}\{\tilde{\Theta}^{(k)}\})\right\}.\end{aligned} \tag{6.31}$$

where we have introduced the associated dual variables. By regrouping, we can write

$$\mathcal{L} = \quad s\left(1 - \mathrm{Tr}\left\{X(\mathbf{id}\otimes\Theta^{(n)})\right\} - \textstyle\sum_{k=2}^{n}\mathrm{Tr}\left\{y^{(k)}(\mathbf{id}_{(2k-3)}\otimes\Theta^{(k-1)} - \mathbf{id}_{(2k-2)}\otimes\Theta^{(k)})\right\}\right) \tag{6.32}$$

$$+\mathrm{Tr}\left\{\textstyle\sum_{a,\alpha}\omega_{a\alpha}\mathsf{T}_{a|\alpha}\right\} + \mathrm{Tr}\left\{\textstyle\sum_{\vec{a}}\left[X - \sum_{a,\alpha}\omega_{a\alpha}d_{\vec{a}}(a|\alpha) - x_{\vec{a}}\right]\tilde{\mathsf{Q}}_{\vec{a}}\right\}. \tag{6.33}$$

Dual SDP

The Lagrangian is rendered independent of the primal variables when $\mathrm{Tr}[X(\mathbf{id}\otimes\Theta^{(n)})]+\sum_{k=2}^{n}\mathrm{Tr}[y^{(k)}(\mathbf{id}_{(2k-3)}\otimes\Theta^{(k-1)}-\mathbf{id}_{(2k-2)}\otimes\Theta^{(k)})] = 1$, $X = \sum_{a,\alpha}\omega_{a\alpha}d_{\vec{a}}(a|\alpha)+$

$x_{\vec{a}}\ \forall\vec{a}$. As such, we can express the dual SDP as,

$$1+\mathcal{R}(\{\mathsf{T}_\alpha\}) = \max_{y,X,\{\omega_{a\alpha}\}} \mathrm{Tr}\left\{\sum_{a,\alpha}\omega_{a\alpha}\mathsf{T}_{a|\alpha}\right\} \tag{6.34}$$

$$\text{s.t. } X \geq \sum_{a,\alpha}\omega_{a\alpha}d_{\vec{a}}(a|\alpha),\ \ \omega_{a\alpha}\geq 0$$

$$\mathrm{Tr}[X(\mathbf{id}\otimes\Theta^{(n)})] + \sum_{k=2}^{n}\mathrm{Tr}[y^{(k)}(\mathbf{id}_{(2k-3)}\otimes\Theta^{(k-1)} - \mathbf{id}_{(2k-2)}\otimes\Theta^{(k)})] = 1. \tag{6.35}$$

To check feasibility, we can find solutions for $y^{(k)}$s occuring in Eq. (6.35), those satisfy $\mathrm{Tr}[X(\mathbf{id}\otimes\Theta^{(n)})]\leq 1$. We find that it is satisfied when we write $\mathrm{Tr}[X]\leq 1/\bar{D}$, where $D=\prod_{j=0}^{n} d_{2n-(2j+1)}$ and d_x is the dimension of the system labeled "x". This translate into the following final form of the SDP,

$$1+\mathcal{R}(\{\mathsf{T}_\alpha\}) = \max_{X,\{\omega_{a\alpha}\}} \mathrm{Tr}\left\{\sum_{a,\alpha}\omega_{a\alpha}\mathsf{T}_{a|\alpha}\right\} \tag{6.36}$$

$$\text{s.t. } X \geq \sum_{a,\alpha}\omega_{a\alpha}d_{\vec{a}}(a|\alpha)$$

$$\omega_{a\alpha}\geq 0,\ \ \mathrm{Tr}[X]\leq 1/\bar{D}.$$

Equivalence between the primal and dual SDP occurs when strong duality exists, when a strictly feasible solution exists for the dual problem. We can manually find that one such solution is $X=(\mathbf{id}_{2n-1}\otimes\mathbf{id})/(D\bar{D})$, and $\omega_{a\alpha}=\gamma\mathbf{id}_{2n-1}\otimes\mathbf{id}$ for γ such that $1/(mD\bar{D})>\gamma>0$ for $m\in\mathbb{R}^+$, where $\bar{D}=\prod_{j=1}^{2n} d_{2n-j}$.

Lower bound

Considering the optimal dual variables y^*, and $\omega^*_{a\alpha}$, we have

$$1+\mathcal{R}(\{\mathsf{T}_\alpha\}) = \mathrm{Tr}\left\{\sum_{a,\alpha}\omega^*_{a\alpha}\mathsf{T}_{a|\alpha}\right\} \tag{6.37}$$

$$\text{s.t. } X^* \geq \sum_{a,\alpha}\omega^*_{a\alpha}d_{\vec{a}}(a|\alpha)$$

$$\omega^*_{a\alpha}\geq 0,\ \ \mathrm{Tr}[X^*]\leq 1/\bar{D}.$$

To eventually draw a connection to the game, let us introduce three auxiliary variables,

$$\ell^* = \mathrm{Tr}\left\{\sum_{a,\alpha}\omega^*_{a\alpha}\right\},\ \ q^*(a,\alpha)=\frac{\mathrm{Tr}\{\omega^*_{a\alpha}\}}{\ell^*}\ \ \text{and}\ \ \mathsf{C}^*_{a|\alpha}=\frac{\omega^*_{a\alpha}}{\ell^* q^*(a,\alpha)}, \tag{6.38}$$

where $\{q^*(a,\alpha)\}$ is normalized probability distributions and ℓ^* is introduced such that $\mathrm{Tr}\left\{\mathsf{C}^*_{a|\alpha}\right\} = D$. Then, Eq.(6.37), can be rewritten as

$$1 + \mathcal{R}(\{\mathsf{T}_\alpha\}) = \ell^* \sum_{a,\alpha} q^*(a,\alpha)\mathrm{Tr}\left\{\mathsf{C}^*_{a|\alpha}\mathsf{T}_{a|\alpha}\right\} \tag{6.39}$$

Let us consider the following choice of strategy by Alice. After receiving the information regarding the ensemble of channels $\{\mathcal{X}^*_\beta\}_\beta$ with $\mathcal{X}^*_\beta = \{\mathcal{C}^*_{b|\beta}, w^*(b|\beta)\}_b$, she plays the game with $p(v) = \delta_{v,0}$ (i.e., sets $v = 0$), $p(\alpha|\beta, v = 0) = \delta_{\alpha,\beta}$ (i.e., applies T_β given β) and $p(g|a,\beta, v = 0) = \delta_{g,b}$ (i.e., guesses $b = a = g$ after getting outcome a). Choosing this sub-optimal strategies, one gets,

$$\begin{aligned} p(\{\mathcal{X}^*_\beta\}, \{\mathsf{T}_\alpha\}) &\geq \sum q^*(b,\beta)\delta_{\lambda,0}\delta_{\alpha,\beta}\mathrm{Tr}\left\{\mathsf{C}^*_{b|\beta}\mathsf{T}_{a|\alpha}\right\}\delta_{g,b}\delta_{a,g}, \\ &= \sum_{a,\alpha} q^*(a,\alpha)\mathrm{Tr}\left\{\mathsf{C}^*_{a|\alpha}\mathsf{T}_{a|\alpha}\right\} \\ &= (1/\ell^*)[1 + \mathcal{R}(\{\mathsf{T}_\alpha\})]. \end{aligned} \tag{6.40}$$

The above SDP constrain so as to furnish the following relation concerning the chosen strategy,

$$X^* \geq \sum_{b,\beta} d_{\vec{b}}(b|\beta)\ell^* q^*(b,\beta)\mathsf{C}^*_{b|\beta}. \tag{6.41}$$

We multiply both sides by an arbitrary but appropriate tester $\tilde{Q}_{\vec{b}}$, sum over $\vec{b}$ and take trace, we are left with,

$$\mathrm{Tr}\left\{X^* \sum_{\vec{b}} \tilde{Q}_{\vec{b}}\right\} \geq \sum_{b,\beta,\vec{b}} d_{\vec{b}}(g|\beta)\delta_{b,g}\ell^* q^*(b,\beta)\mathrm{Tr}\left\{\mathsf{C}^*_{b|\beta}\tilde{Q}_{\vec{b}}\right\}. \tag{6.42}$$

Due to normalization of an n-slot tester, the left hand side is equal to one. Moreover, the above equation holds even when we maximize over all $\tilde{Q}_{\vec{b}}$. This leads to,

$$\frac{1}{\ell^*} \geq \max_{\tilde{Q}_{\vec{b}}} \sum_{b,\beta,\vec{b}} q^*(b,\beta)\mathrm{Tr}\left\{\mathsf{C}^*_{b|\beta}\tilde{Q}_{\vec{b}}\right\} d_{\vec{b}}(g|\beta)\delta_{b,g}. \tag{6.43}$$

We observe the resemblance of right hand side of above equation with the success probability in the game with a single tester gameplay. We also note that $(1/\ell^*) \geq p(\{\mathcal{X}^*_\beta\})$. This leads to the following lower bound.

$$1 + \mathcal{R}(\{\mathsf{T}_\alpha\}) \leq \frac{p(\{\mathcal{X}^*_\beta\}, \{\mathsf{T}_\alpha\})}{p(\{\mathcal{X}^*_\beta\})}. \tag{6.44}$$

Since $\max_{\{\mathcal{X}_\beta\}} \frac{p(\{\mathcal{X}_\beta\},\{\mathsf{T}_\alpha\})}{p(\{\mathcal{X}_\beta\})}$ is bounded above and below by the same quantity, we can conclude that it is equal to the quantity $1 + \mathcal{R}(\{\mathsf{T}_\alpha\})$, thus establishing our result,

$$\max_{\{\mathcal{X}_\beta\}} \frac{p(\{\mathcal{X}_\beta\}, \{\mathsf{T}_\alpha\})}{p(\{\mathcal{X}_\beta\})} = 1 + \mathcal{R}(\{\mathsf{T}_\alpha\}). \tag{6.45}$$

6.4 Ramifications

As innocent as the theorem revolving around Theorem 6.2.1 might present itself, it has wider implications because of the general nature of quantum combs and testers in describing all possible quantum devices, from states and single time-step channels to incomplete quantum networks. In the following we discuss some ramifications of our finding.

- ***On state distinguishability.*** A special case of our comb distinguishability game is when all the combs are quantum states. Consequently, this reduces the involved testers to observables and notion of incompatibility reduces to observable incompatibility. As such, the result rereads itself as that all incompatible collections of observables fare better than compatible collections, in the associated state distinguishability game. This is exactly the result found by Skrzypczyk et.al. in [SŠC19]. In this context, our result generalises the one due to Skrzypczyk et.al. as well.
- ***On single time-step quantum channel distinguishability.*** When the game is considered for single time-step quantum channels, we look at a large class of channels which has been investigated from a point of view of distinguishability. One major difference on distinguishability between states and channels is that states are quantum devices of static nature and channels of dynamic, even though all distinguishability tasks boil down to distinguishing states. In the wake of this, we can conclude that the statistical notion of incompatibility showcases its behaviour in its generality.
- ***On memory channel distinguishability.*** Quantum combs describe equivalent classes of memory channels. The higher the order a particular comb belongs to, the higher the number of time-steps or memories associated with the underlying memory channel. Our result establishes that irrespective of the order an ensemble of comb belong to, incompatible testers from that specific order always outperform compatible ones in distinguishing the ensemble. This statistical notion of higher-order incompatibility is surpassing any other possible quantum resources present, including memories. Despite, the notion being statistical in nature, it potentially arises from all of the building components of the underlying causal network of the testers.
- ***On distinguishability with multiple queries.*** Distinguishing states with multiple copies, quantum processes with multiple queries or quantum observables with multiple implementations; all of these multiple objects

are collectively described by memoryless combs. As such, we can always assert that irrespective whether you have access to multiple shots or not, "incompatibility will always outperform compatibility".

Chapter 7

Conclusions

In this thesis, and through the associated research papers, we have introduced the quantum information processing task of observable labeling. Several of the associated investigations of this task have been studied using the framework of higher-order quantum theory. Moreover, pursuing a different direction of study, we found out that higher-order incompatibility is advantageous in quantum comb discrimination games. The concluding remarks on the discoveries related to these two explored avenues are elaborated below.

On quantum labeling tasks

Simultaneous with formalising the task of labeling, we identified that such tasks constitute a specific class of distinguishability tasks concerning quantum observables. As such, these tasks hold relevance when one has access to finite number of implementations of the given unlabeled measurement devices, which is when they do not have access to complete statistics from the experiments. Consequently, we investigated these distinguishability tasks in the single-shot as well as the multiple-shot regimes.

In the single-shot regime, we progressed our investigation from the case of binary observables to non-binary. We discovered that only[1] rank-deficient binary observables can be perfectly labeled with a single use of the measurement device. Moreover, since our investigation was carried out by considering the most general quantum networks designed to label observables, we found that assistance of entanglement is unnecessary. Alongside, it was also found that full rank binary observables, those which do not admit perfect labeling, neither admit themeselves to be labeled unambiguously by admitting some failed decisions. Then, by evaluating the averaged minimum-error probability associated with labeling binary observables, we identified an operational distance, between the underlying

[1]Not including the pathological cases discussed in the chapter 4 and chapter 5.

effects, that reflects the labelability of the associated unlabeled observable. Moving onto non-binary observables, as one would intuitively expect, we rigorously showed that such observables do not admit any perfect labeling with a single shot. Here also, while studying the minimum-error labeling, we discovered that entanglement does not improve labelability. Winding up the investigations on single shot labeling, we introduced schemes for identifying restricted number of outcome-label associations.

The multiple-shot regime was investigated with the prospect of witnessing improvement in labelability, with the aid of more number of available implementation of the observables; Availability of more number of shots translates to more complex quantum networks, possessing potential quantum resources including quantum memories, adaptivity and entanglement. To our dismay, we discovered that the performance improvement to be not in the affirmative; Binary observbles which were not perfectly labelable, in a single use, were still found to be perfectly labelable with more uses. More crucially, we discovered that, for minimum-error labeling of binary observables, access to multiple uses is counter-intuitively detrimental. The class of observables those benefits from this regime is that of non-binary ones. For them we identified the minimal number uses required for perfectly labeling them and give schemes for partially identifying conducive label-effect associations. To sum up the studies into short capsules, we can have the following ones:

- *Full-rank quantum observables can never be labeled perfectly or unambiguously.* This statement holds in its full generality; Irrespective of the number of uses and potential resources from the most general labeling experiments.
- *Neither assistance due to entanglement, adaptibility nor quantum memory effects improve performance of labeling experiments.*

The task of labeling not only reveal restrictions imposed by quantum theory and consequently be of foundational interest. By construction, it referes to experimental scenarios where the label identities are lost or where the experimenter is doubting the description of the quantum detectors. These scenarios may seem to be highly unlikely and constrived. But, in typical quantum experiments more complex setups can be designed, those which are composed of many measurement devices and detectors, leading to more abstract outcomes. In such scenarios, one can certify the associations of the outcome labels and the effects. This says, in other words, that the risk of loosing the labels and trying to identify them is not a physical problem as such but originates due to the deficiency of the end-users.

Moreover, the investigations regarding labeling have been heavily done using the machinerey of higher-order quantum theoretic notions. This enabled us to come to conclusions which are general in their tones. As such, this avenue pursued in the thesis showcases the potential of higher-order quantum theory as a tool in probing the limits of quantum theory.

On advantageous higher-order incompatibility

In this avenue, we investigated the quality of resourcefulness of incompatible testers in quantum comb distinguishability. We discovered that all collections of incompatible testers result in better performance, when compared to any compatible collection, in a specific quantum comb distinguishability game. Nonetheless, this reveals the resourcefulness of higher-order incompatibility. Our finding generalises an earlier one on '*Quantum observable incompatibility is a resource for state distinguishability*' [SŠC19] to '*Quantum tester incompatibility is a resource for comb distinguishability*'. A major difference of the later statement from the earlier is that the previous one is related to distinguishability of quantum devices of static nature, that are, quantum states and the associated resourcefulness of incompatible quantum testing procedures for states, modeled by quantum observables. The later, our statement, relates to the distinguishability of quantum combs, which are quantum devices primarily of dynamical nature, and incompatibility associated with testing procedures designed to test combs. Nevertheless our result contains the previous one since, states can be considered as special cases of quantum combs and observables as that of quantum testers. Moreover, since quantum combs describe a wide array of quantum devices, from states, single time step quantum evolutions to multiple-time step evolutions with memory, our result is all inclusive in this sense. Another speciality of our finding is that it includes all distinguishability schemes possible for any basic quantum device (including states, single time-step evolutions and POVMs): irrespective of whether it is single-shot or multiple-shot and whether it is parallel or adaptive. In conclusion, with all possible ramifications, we have discovered that the more complex notion of higher-order incompatibility to a resource in quantum hypothesis-testing tasks involving combs.

Bibliography

[ACMnT$^+$07] K. M. R. Audenaert, J. Calsamiglia, R. Muñoz Tapia, E. Bagan, Ll. Masanes, A. Acin, and F. Verstraete. Discriminating states: The quantum chernoff bound. *Phys. Rev. Lett.*, 98:160501, Apr 2007.

[Bae13] Joonwoo Bae. Structure of minimum-error quantum state discrimination. *New Journal of Physics*, 15(7):073037, 2013.

[BCD$^+$09a] A. Bisio, G. Chiribella, G. M. D'Ariano, S. Facchini, and P. Perinotti. Optimal quantum tomography of states, measurements, and transformations. *Phys. Rev. Lett.*, 102:010404, Jan 2009.

[BCD$^+$09b] A. Bisio, G. Chiribella, G. M. D'Ariano, S. Facchini, and P. Perinotti. Optimal quantum tomography of states, measurements, and transformations. *Phys. Rev. Lett.*, 102:010404, Jan 2009.

[BCW$^+$12] Cyril Branciard, Eric G Cavalcanti, Stephen P Walborn, Valerio Scarani, and Howard M Wiseman. One-sided device-independent quantum key distribution: Security, feasibility, and the connection with steering. *Physical Review A*, 85(1):010301, 2012.

[BH13] Joonwoo Bae and Won-Young Hwang. Minimum-error discrimination of qubit states: Methods, solutions, and properties. *Phys. Rev. A*, 87:012334, Jan 2013.

[BJOP14] Somshubhro Bandyopadhyay, Rahul Jain, Jonathan Oppenheim, and Christopher Perry. Conclusive exclusion of quantum states. *Physical Review A*, 89(2):022336, 2014.

[BLBSM23] Jessica Bavaresco, Patryk Lipka-Bartosik, Pavel Sekatski, and Mohammad Mehboudi. Designing optimal protocols in bayesian quantum parameter estimation with higher-order operations. *arXiv preprint arXiv:2311.01513*, 2023.

[BLPY16] Paul Busch, Pekka Lahti, Juha-Pekka Pellonpää, and Kari Ylinen. *Quantum measurement*, volume 23. Springer, 2016.

[BMQ21] Jessica Bavaresco, Mio Murao, and Marco Túlio Quintino. Strict hierarchy between parallel, sequential, and indefinite-causal-order strategies for channel discrimination. *Physical review letters*, 127(20):200504, 2021.

[BMQ22] Jessica Bavaresco, Mio Murao, and Marco Túlio Quintino. Unitary channel discrimination beyond group structures: Advantages of sequential and indefinite-causal-order strategies. *Journal of Mathematical Physics*, 63(4), 2022.

[BNV13] Nicolas Brunner, Miguel Navascués, and Tamás Vértesi. Dimension witnesses and quantum state discrimination. *Physical review letters*, 110(15):150501, 2013.

[BPA+08] Nicolas Brunner, Stefano Pironio, Antonio Acin, Nicolas Gisin, André Allan Méthot, and Valerio Scarani. Testing the dimension of hilbert spaces. *Physical review letters*, 100(21):210503, 2008.

[CDP08a] Giulio Chiribella, G Mauro D'Ariano, and Paolo Perinotti. Transforming quantum operations: Quantum supermaps. *Europhysics Letters*, 83(3):30004, 2008.

[CDP08b] Giulio Chiribella, Giacomo Mauro D'Ariano, and Paolo Perinotti. Optimal cloning of unitary transformation. *Phys. Rev. Lett.*, 101:180504, Oct 2008.

[CDP08c] Giulio Chiribella, G Mauro D'Ariano, and Paolo Perinotti. Quantum circuit architecture. *Physical review letters*, 101(6):060401, 2008.

[CDP08d] Giulio Chiribella, Giacomo M D'Ariano, and Paolo Perinotti. Memory effects in quantum channel discrimination. *Physical review letters*, 101(18):180501, 2008.

[CDP09] Giulio Chiribella, Giacomo Mauro D'Ariano, and Paolo Perinotti. Theoretical framework for quantum networks. *Physical Review A*, 80(2):022339, 2009.

[CFS02] Carlton M Caves, Christopher A Fuchs, and Rüdiger Schack. Conditions for compatibility of quantum-state assignments. *Physical Review A*, 66(6):062111, 2002.

[Chi12] Giulio Chiribella. Perfect discrimination of no-signalling channels via quantum superposition of causal structures. *Phys. Rev. A*, 86:040301, Oct 2012.

[Cho75] Man-Duen Choi. Completely positive linear maps on complex matrices. *Linear algebra and its applications*, 10(3):285–290, 1975.

[CHT19] Claudio Carmeli, Teiko Heinosaari, and Alessandro Toigo. Quantum incompatibility witnesses. *Phys. Rev. Lett.*, 122:130402, Apr 2019.

[CMnTM+08] J. Calsamiglia, R. Muñoz Tapia, Ll. Masanes, A. Acin, and E. Bagan. Quantum chernoff bound as a measure of distinguishability between density matrices: Application to qubit and gaussian states. *Phys. Rev. A*, 77:032311, Mar 2008.

[DBSA21] Chandan Datta, Tanmoy Biswas, Debashis Saha, and Remigiusz Augusiak. Perfect discrimination of quantum measurements using entangled systems. *New Journal of Physics*, 23(4):043021, 2021.

[DL70] E Brian Davies and John T Lewis. An operational approach to quantum probability. *Communications in Mathematical Physics*, 17(3):239–260, 1970.

[DPS11] Giacomo Mauro D'Ariano, Paolo Perinotti, and Michal Sedlák. Extremal quantum protocols. *Journal of mathematical physics*, 52(8), 2011.

[Fin82] Arthur Fine. Hidden variables, joint probability, and the bell inequalities. *Physical Review Letters*, 48(5):291, 1982.

[GW07] Gus Gutoski and John Watrous. Toward a general theory of quantum games. In *Proceedings of the Thirty-Ninth Annual ACM Symposium on Theory of Computing*, STOC '07, page 565–574, New York, NY, USA, 2007. Association for Computing Machinery.

[GZ21] Seyed Arash Ghoreishi and Mario Ziman. Minimum-error discrimination of thermal states. *Physical Review A*, 104(6):062402, 2021.

[Hei05] Teiko Heinonen. Optimal measurements in quantum mechanics. *Physics Letters A*, 346(1-3):77–86, 2005.

[Hel69] Carl W Helstrom. Quantum detection and estimation theory. *Journal of Statistical Physics*, 1:231–252, 1969.

[HHLW10] Aram W. Harrow, Avinatan Hassidim, Debbie W. Leung, and John Watrous. Adaptive versus nonadaptive strategies for quantum channel discrimination. *Phys. Rev. A*, 81:032339, Mar 2010.

[HHP12] Erkka Haapasalo, Teiko Heinosaari, and Juha-Pekka Pellonpää. Quantum measurements on finite dimensional systems: relabeling and mixing. *Quantum Information Processing*, 11:1751–1763, 2012.

[HK18] Teiko Heinosaari and Oskari Kerppo. Antidistinguishability of pure quantum states. *Journal of Physics A: Mathematical and Theoretical*, 51(36):365303, 2018.

[HM17] Teiko Heinosaari and Takayuki Miyadera. Incompatibility of quantum channels. *Journal of Physics A: Mathematical and Theoretical*, 50(13):135302, 2017.

[HMZ16] Teiko Heinosaari, Takayuki Miyadera, and Mário Ziman. An invitation to quantum incompatibility. *Journal of Physics A: Mathematical and Theoretical*, 49(12):123001, 2016.

[HN03] Masahito Hayashi and Hiroshi Nagaoka. General formulas for capacity of classical-quantum channels. *IEEE Transactions on Information Theory*, 49(7):1753–1768, 2003.

[Hol98] Alexander S Holevo. Coding theorems for quantum channels. *arXiv preprint quant-ph/9809023*, 1998.

[Hol11] Alexander S Holevo. *Probabilistic and statistical aspects of quantum theory*, volume 1. Springer Science & Business Media, 2011.

[HRRZ18] Teiko Heinosaari, Daniel Reitzner, Tomáš Rybár, and Mário Ziman. Incompatibility of unbiased qubit observables and pauli channels. *Physical Review A*, 97(2):022112, 2018.

[HZ11] Teiko Heinosaari and Mário Ziman. *The mathematical language of quantum theory: from uncertainty to entanglement*. Cambridge University Press, 2011.

[Jam72] Andrzej Jamiołkowski. Linear transformations which preserve trace and positive semidefiniteness of operators. *Reports on Mathematical Physics*, 3(4):275–278, 1972.

[JC24] Kaiyuan Ji and Eric Chitambar. Incompatibility as a resource for programmable quantum instruments. *PRX Quantum*, 5(1):010340, 2024.

[Jen12a] Anna Jenčová. Extremality conditions for generalized channels. *Journal of mathematical physics*, 53(12), 2012.

[Jen12b] Anna Jenčová. Generalized channels: channels for convex subsets of the state space. *Journal of mathematical physics*, 53(1), 2012.

[Jen13] Anna Jenčová. Extremal generalized quantum measurements. *Linear Algebra and its Applications*, 439(12):4070–4079, 2013.

[Jen21] Anna Jencova. A general theory of comparison of quantum channels (and beyond). *IEEE Transactions on Information Theory*, 67(6):3945–3964, 2021.

[JFDY06] Zhengfeng Ji, Yuan Feng, Runyao Duan, and Mingsheng Ying. Identification and distance measures of measurement apparatus. *Physical Review Letters*, 96(20):200401, 2006.

[KDM+24] Stanislaw Kurdzialek, Piotr Dulian, Joanna Majsak, Sagnik Chakraborty, and Rafal Demkowicz-Dobrzanski. Quantum metrology using quantum combs and tensor network formalism. *arXiv preprint arXiv:2403.04854*, 2024.

[KM24] Taihei Kimoto and Takayuki Miyadera. Upper bounds on probabilities in channel measurements on qubit channels and their applications, 2024.

[KPP20] Aleksandra Krawiec, Łukasz Pawela, and Zbigniew Puchała. Discrimination of povms with rank-one effects. *Quantum Information Processing*, 19:1–12, 2020.

[KW05] Dennis Kretschmann and Reinhard F Werner. Quantum channels with memory. *Physical Review A*, 72(6):062323, 2005.

[Lah03] Pekka Lahti. Coexistence and joint measurability in quantum mechanics. *International Journal of Theoretical Physics*, 42:893–906, 2003.

[LHYY23] Qiushi Liu, Zihao Hu, Haidong Yuan, and Yuxiang Yang. Optimal strategies of quantum metrology with a strict hierarchy. *Phys. Rev. Lett.*, 130:070803, Feb 2023.

[LP97] Pekka Lahti and Sylvia Pulmannová. Coexistent observables and effects in quantum mechanics. *Reports on Mathematical Physics*, 39(3):339–351, 1997.

[LS24] Leevi Leppäjärvi and Michal Sedlák. Incompatibility of quantum instruments. *Quantum*, 8:1246, 2024.

[MAG06] Ll Masanes, Antonio Acín, and Nicolas Gisin. General properties of nonsignaling theories. *Physical Review A*, 73(1):012112, 2006.

[MdM90] Hans Martens and Willem M de Muynck. Nonideal quantum measurements. *Foundations of Physics*, 20(3):255–281, 1990.

[MF22] Arindam Mitra and Máté Farkas. Compatibility of quantum instruments. *Physical Review A*, 105(5):052202, 2022.

[OCB12] Ognyan Oreshkov, Fabio Costa, and Časlav Brukner. Quantum correlations with no causal order. *Nature communications*, 3(1):1092, 2012.

[OKSM24] Tatsuki Odake, Hlér Kristjánsson, Akihito Soeda, and Mio Murao. Higher-order quantum transformations of hamiltonian dynamics. *Physical Review Research*, 6(1):L012063, 2024.

[Pau02] Vern Paulsen. *Completely bounded maps and operator algebras*, volume 78. Cambridge University Press, 2002.

[Per97] Asher Peres. *Quantum theory: concepts and methods*, volume 72. Springer, 1997.

[PPK+21] Zbigniew Puchała, Łukasz Pawela, Aleksandra Krawiec, Ryszard Kukulski, and Michał Oszmaniec. Multiple-shot and unambiguous discrimination of von neumann measurements. *Quantum*, 5:425, 2021.

[PPKK18] Zbigniew Puchała, Łukasz Pawela, Aleksandra Krawiec, and Ryszard Kukulski. Strategies for optimal single-shot discrimination of quantum measurements. *Physical Review A*, 98(4):042103, 2018.

[PZ22] Jaroslav Pavličko and Mário Ziman. Robustness of optimal probabilistic storage and retrieval of unitary channels to noise. *Physical Review A*, 106(5):052416, 2022.

[QVB14] Marco Túlio Quintino, Tamás Vértesi, and Nicolas Brunner. Joint measurability, einstein-podolsky-rosen steering, and bell nonlocality. *Physical review letters*, 113(16):160402, 2014.

[Rag02] Maxim Raginsky. Strictly contractive quantum channels and physically realizable quantum computers. *Physical Review A*, 65(3):032306, 2002.

[Sac05] Massimiliano F Sacchi. Optimal discrimination of quantum operations. *Physical Review A*, 71(6):062340, 2005.

[SBZ19] Michal Sedlák, Alessandro Bisio, and Mário Ziman. Optimal probabilistic storage and retrieval of unitary channels. *Physical review letters*, 122(17):170502, 2019.

[SC23] Paul Skrzypczyk and Daniel Cavalcanti. *Semidefinite Programming in Quantum Information Science*. IOP Publishing, 2023.

[Sed10] Michal Sedlák. Quantum theory of unambiguous measurements. *arXiv preprint arXiv:1003.2448*, 2010.

[SRCZ16] Michal Sedlák, Daniel Reitzner, Giulio Chiribella, and Mário Ziman. Incompatible measurements on quantum causal networks. *Physical Review A*, 93(5):052323, 2016.

[SŠC19] Paul Skrzypczyk, Ivan Šupić, and Daniel Cavalcanti. All sets of incompatible measurements give an advantage in quantum state discrimination. *Physical review letters*, 122(13):130403, 2019.

[SZ09] Michal Sedlák and Mário Ziman. Unambiguous comparison of unitary channels. *Physical Review A*, 79(1):012303, 2009.

[SZ14] Michal Sedlák and Mário Ziman. Optimal single-shot strategies for discrimination of quantum measurements. *Physical Review A*, 90(5):052312, 2014.

[UMG14] Roope Uola, Tobias Moroder, and Otfried Gühne. Joint measurability of generalized measurements implies classicality. *Physical review letters*, 113(16):160403, 2014.

[vNB18] John von Neumann and ROBERT T. BEYER. *Mathematical Foundations of Quantum Mechanics: New Edition*. Princeton University Press, ned - new edition, 2018.

[WPGF09] Michael M. Wolf, David Perez-Garcia, and Carlos Fernandez. Measurements incompatible in quantum theory cannot be measured jointly in any other no-signaling theory. *Phys. Rev. Lett.*, 103:230402, Dec 2009.

[Yan19] Yuxiang Yang. Memory effects in quantum metrology. *Physical review letters*, 123(11):110501, 2019.

[YSM23] Satoshi Yoshida, Akihito Soeda, and Mio Murao. Reversing unknown qubit-unitary operation, deterministically and exactly. *Physical Review Letters*, 131(12):120602, 2023.

[ZH08] Mario Ziman and Teiko Heinosaari. Discrimination of quantum observables using limited resources. *Physical Review A*, 77(4):042321, 2008.

[ZHS09] Mario Ziman, Teiko Heinosaari, and Michal Sedlák. Unambiguous comparison of quantum measurements. *Physical Review A*, 80(5):052102, 2009.

[Zim08] Mário Ziman. Process positive-operator-valued measure: A mathematical framework for the description of process tomography experiments. *Physical Review A*, 77(6):062112, 2008.

[ZS10] Mário Ziman and Michal Sedlák. Single-shot discrimination of quantum unitary processes. *Journal of Modern Optics*, 57(3):253–259, 2010.