

# An estimation theoretic approach to quantum realizability problems

by

Thomas C. Fraser

A thesis  
presented to the University of Waterloo  
in fulfillment of the  
thesis requirement for the degree of  
Doctor of Philosophy  
in  
Physics

Waterloo, Ontario, Canada, 2023

© Thomas C. Fraser 2023

## Examining Committee Membership

The following served on the Examining Committee for this thesis. The decision of the Examining Committee is by majority vote.

External  
Examiner: Renato Renner  
Professor, Dept. of Physics,  
ETH Zürich

Supervisor: Robert Spekkens  
Research Faculty,  
Perimeter Institute for Theoretical Physics

Co-supervisor: Kevin Resch  
Professor, Dept. of Physics & Astronomy,  
University of Waterloo

Internal Member: Thomas Jennewein  
Associate Professor, Dept. of Physics & Astronomy,  
University of Waterloo

Internal-External  
Member: William Slofstra  
Associate Professor, Dept. of Pure Mathematics,  
University of Waterloo

Other Member: Lucien Hardy  
Research Faculty,  
Perimeter Institute for Theoretical Physics

## **Author's Declaration**

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

## Abstract

This thesis seeks to develop a general method for solving so-called *quantum realizability problems*, which are questions of the following form: under which conditions does there exist a quantum state exhibiting a given collection of properties? The approach adopted by this thesis is to utilize mathematical techniques previously developed for the related problem of *property estimation* which is concerned with learning or estimating the properties of an *unknown* quantum state. Our primary result is to recognize a correspondence between (i) property values which are *realized* by some quantum state, and (ii) property values which are *occasionally* produced as estimates of a generic quantum state.

In [Chapter 3](#), we review the concepts of stability and norm minimization from geometric invariant theory and non-commutative optimization theory for the purposes of characterizing the flow of a quantum state under the action of a reductive group. In particular, we discover that most properties of quantum states are related to the gradient of this flow, also known as the *moment map*. Afterwards, [Chapter 4](#) demonstrates how to *estimate* the value of the moment map of a quantum state by performing a covariant quantum measurement on a large number of identical copies of the quantum state. These measurement schemes for estimating the moment map of a quantum state arise naturally from the decomposition of a large tensor-power representation into its irreducible sub-representations. Then, in [Chapter 5](#), we prove an exact correspondence between the realizability of a moment map value on one hand and the asymptotic likelihood it is produced as an estimate on the other hand. In particular, by composing these estimation schemes, we derive necessary and sufficient conditions for the existence of a quantum state jointly realizing any finite collection of moment maps.

Finally, in [Chapter 6](#) we apply these techniques to the *quantum marginals problem* which aims to characterize precisely the relationships between the marginal density operators describing the various subsystems of a composite quantum system. We make progress toward an analytic solution to the quantum marginals problem by deriving a complete hierarchy of necessary inequality constraints.

## Acknowledgements

First I must thank my supervisor Robert Spekkens for his sustained support and patience. Without reservation, Rob generously shared with me his time, knowledge and insight, and for that I am truly indebted. I am thankful for Rob's courage in gifting me with the freedom to develop my own intuitions, pursue my own research ambitions and thus grow as an independent researcher. It has been a privilege to work at the Perimeter Institute in Rob's research group.

Over the course of my time as a student of mathematics and physics, I had the tremendous pleasure to meet and share ideas with a bounty of bright individuals. Foremost among them is my colleague and dear friend Jack Davis. The discussions and adventures I have shared with Jack have been invigorating and memorable. His influence on my approach to physics and life is immeasurable. I also must express my gratitude to Tomáš Gonda for his spiritual and intellectual guidance, his appetite for abstraction and his tea parties. I am thankful for Elie Wolfe's unwavering encouragement and belief in me. I am grateful to John Selby for showing me how to think and calculate diagrammatically, and more generally, for teaching me the active role played by notation.

In addition to those already mentioned, I am thankful for my interactions with, in no particular order, Daniel Grimmer, David Schmid, Tobias Fritz, Thomas Galley, Jake Ferguson, Finnian Gray, Denis Rosset, William Slofstra, Benjamin Lovitz, Ding Jia, Jamie Sikora, Nitica Sakharwalde, Robin Lorenz, Pedro Lauand, Marina Ansanelli, Matthew Fox, Isaac Smith, Flaminia Giacomini, Albert Werner, Freek Witteveen, Mark Wilde, Julia Liebert, Lexin Ding, Ravi Kunjwal and Belén Sainz.

I am grateful to Matthias Christandl, Christian Schilling, and Gilad Gour for the privilege of allowing me to visit their respective research groups and receive a wealth of new ideas and fresh perspectives.

At the Perimeter Institute, where I spent some of my time as an undergraduate student and all of my time as a graduate student, I was supported by countless individuals including Debbie Guenther, Jamie Foley, and everyone in the Black hole bistro. Finally, I must thank both the members of my PhD advisory committee for their feedback and guidance and the members of my thesis examining committee for their valuable time and expertise.

## Dedication

Without the strength of my mother, the genius of my father, or the love and sense of humor of my sister, I would not be here today. This thesis is dedicated to them.

# Table of Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Quantum estimation theory . . . . .	2
1.2	Quantum realizability problems . . . . .	5
1.3	Organization . . . . .	8
1.4	A toy example . . . . .	9
<b>2</b>	<b>Preliminaries</b>	<b>15</b>
2.1	Measure theory . . . . .	15
2.1.1	Measures . . . . .	15
2.1.2	Weak convergence . . . . .	17
2.1.3	Moments & cumulants . . . . .	21
2.1.4	Rate functions . . . . .	22
2.1.5	Principles of large deviations . . . . .	22
2.2	Quantum theory . . . . .	24
2.2.1	Quantum states . . . . .	24
2.2.2	Measurements . . . . .	29
2.3	Group theory . . . . .	30
2.3.1	Groups . . . . .	30
2.3.2	Lie groups . . . . .	32
2.3.3	Lie algebras . . . . .	33
2.4	Representation theory . . . . .	35
2.4.1	Group representations . . . . .	35
2.4.2	Lie algebra representations . . . . .	36
2.4.3	Composition . . . . .	39
2.4.4	Decomposition . . . . .	41
2.4.5	Complexification & semisimplicity . . . . .	42
2.4.6	Roots systems . . . . .	48
2.4.7	Highest weights . . . . .	50
2.5	Glossary of Notation . . . . .	57
<b>3</b>	<b>Non-commutative optimization</b>	<b>59</b>
3.1	Introduction . . . . .	59
3.2	Invariants & norm minimization . . . . .	60
3.2.1	Group orbits & stability . . . . .	60
3.2.2	Invariant and fixed subspaces . . . . .	64

3.2.3	Capacities & moment maps . . . . .	66
3.2.4	Composition of capacities & moment maps . . . . .	71
3.3	Occasionality & semistability . . . . .	72
3.3.1	Cumulants of quantum observables . . . . .	72
3.3.2	Typical, occasional & exceptional behaviours . . . . .	75
3.3.3	The occasionality theorem . . . . .	78
3.3.4	Strong duality . . . . .	83
<b>4</b>	<b>Estimation theory</b> . . . . .	<b>85</b>
4.1	Estimating a probability distribution . . . . .	85
4.2	Quantum measurements & representations . . . . .	89
4.2.1	Irreducible measurements . . . . .	89
4.2.2	Completely reducible measurements . . . . .	91
4.3	Deformed strong duality . . . . .	93
4.3.1	The inversion trick . . . . .	94
4.3.2	Deforming moment maps . . . . .	95
4.3.3	Rotating & scaling extremal weight vectors . . . . .	97
4.3.4	Deformed capacity . . . . .	99
4.3.5	Strong duality & purifications . . . . .	101
4.4	Estimation schemes . . . . .	102
4.4.1	Estimating moment maps . . . . .	102
4.4.2	General definition . . . . .	106
4.4.3	Derived estimation schemes . . . . .	109
4.4.4	Example estimation schemes . . . . .	114
<b>5</b>	<b>Realizability problems</b> . . . . .	<b>117</b>
5.1	Realizability . . . . .	117
5.1.1	Quantitative realizability . . . . .	119
5.1.2	Inclusivity of de Finetti states . . . . .	122
5.1.3	Moment polytopes . . . . .	123
5.1.4	Beyond moment polytopes . . . . .	126
5.1.5	The biriffle formula . . . . .	128
5.2	Examples . . . . .	132
5.2.1	Absolutely maximally entangled states . . . . .	132
5.2.2	Quantum marginals . . . . .	134
<b>6</b>	<b>Quantum marginal problems</b> . . . . .	<b>135</b>
6.1	Introduction . . . . .	135
6.2	A Reformulation of the QMP . . . . .	137
6.3	Necessary and Sufficient Inequality Constraints . . . . .	138
6.4	Conclusion . . . . .	141
6.5	Supporting results . . . . .	141
6.5.1	Schur-Weyl Decompositions . . . . .	141
6.5.2	Spectra & Partitions . . . . .	143
6.5.3	Proof of Lemma 6.3.5 . . . . .	145



6.5.4	Keyl Divergence & State Discrimination . . . . .	145
6.6	Special Cases . . . . .	149
6.6.1	Pure vs. Full QMP . . . . .	149
6.6.2	Diagrammatics . . . . .	150
6.6.3	The degree one case . . . . .	152
6.6.4	The Bipartite QMP . . . . .	154
6.6.5	Fermionic & Bosonic QMP . . . . .	156
6.6.6	Counting Solutions to the QMP . . . . .	156
	<b>Concluding remarks</b>	<b>158</b>
	<b>Bibliography</b>	<b>161</b>
	<b>Appendices</b>	<b>176</b>
A.1	Misc. Results . . . . .	176
A.2	Invariant post-selection . . . . .	176
A.3	Expected expectation values . . . . .	179
A.4	A quantum de Finetti theorem . . . . .	183
A.5	The Laplace principle . . . . .	185
A.6	Asymptotic Born rule . . . . .	187

# Chapter 1

## Introduction

When provided with a description of some physical system, often called a model or a state, together with a specified measurement or experiment to be performed upon that system, the problem of predicting the result of that experiment is known as the *forward problem*. The *inverse problem*, on the other hand, is to calculate or reconstruct, from the results of the experiment, a description of the physical system that was measured. In either case, a fundamental challenge to overcome is the universal fact that descriptions of physical systems must, for both practical and fundamental reasons, be considered incomplete.

Within the context of quantum theory, a quantum state is a mathematical object which encodes information about a system that is deemed sufficient to make predictions about the statistical behaviour of any hypothetical experiment. Nevertheless, there are numerous applications of quantum theory wherein only a fraction of this information is available or actually required. In these situations, it oftens becomes computationally and conceptually useful to derive or construct an effective theory which is merely concerned with the features or properties of the quantum state that are relavent for the particular context, together with a characterization of the relationships or constraints satisfied by those properties.

The purpose of this thesis is to describe a particular strategy for understanding the relationships between the various properties of a quantum state that is based upon insights from the representation theory of groups for the purposes of tomography and property estimation. It will be shown that this technique asymptotically decides whether or not a given collection of property values can be *realized* by any quantum state, and moreover, can be used to approximately determine what *proportion* of quantum states exhibits those property values. Although this approach is *asymptotic* in nature, meaning it only provides an approximate understanding which becomes exact in the appropriate limiting cases, it is rather *universal* in that it applies to a large class of properties that might be of interest.

To begin, [Section 1.1](#) considers the subject of quantum tomography which seeks to learn or estimate the properties of an unknown quantum state by performing a collective measurement on many identical copies of that state. In particular, we emphasize the role played by the representation theory of groups in the construction of quantum measurements whose outcomes correspond to *estimates* of the invariant and covariant properties of the quantum states they are performed on.

In [Section 1.2](#), we turn our attention to the relational point of view which aims to understand how the various properties of a quantum system relate to one another. Our

primary focus is on the question of *realizability* which asks: given a finite collection of properties, which values for those properties can be jointly realized by some quantum state? In [Section 1.2](#), we will briefly review a handful of problems in quantum information theory which can be formulated as questions of this form, which we refer to as *quantum realizability problems*.

The objective of this thesis, in the end, is to develop a method for solving quantum realizability problems by using insights from the theory of quantum tomography. In [Section 1.3](#), the overall structure of the thesis is outlined, along with a brief summary of the contents of each chapter. Finally, [Section 1.4](#) concludes with a demonstration of the central themes of the thesis through the lense of a simple toy example.

## 1.1 Quantum estimation theory

A fundamental task in quantum information theory, referred to as *quantum tomography*, is determination of the state or characteristics of a quantum system by means of repeated experimentation [[PR04](#); [Hol11](#)]. Following the foundational papers of Fano [[Fan57](#)], Helstrom [[Hel69](#)] and Vogel and Risken [[VR89](#)], the general paradigm is to consider the independent preparation of  $n$  identical copies of a quantum state along with a strategically designed measurement procedure whose outcomes can then be used to produce an estimate for either the values of some of its properties [[Hol78](#)], or more generally, the identity of the entire quantum state [[DPS03](#)].

Generally speaking, there exists a myriad of factors one might wish to optimize for in the context of quantum tomography, including various measures of estimation error [[AKG19](#)], the number of copies needed to achieve a certain threshold of accuracy [[MP95](#); [Haa+16](#)], the optimal estimate for fixed finite  $n$  [[MP95](#)], the adaptability of a measurement scheme to previous data [[Str16](#)], the finiteness of measurement outcomes [[DBE98](#)], and/or the asymptotics as  $n$  tends to infinity [[Hay05](#); [GM05](#); [Key06](#)].

Once a measurement protocol has been selected and performed, there are a variety of strategies for converting the obtained measurement data into an estimate for identity of the state that was measured, each of which exhibits its own advantages and disadvantages. If the performed measurements are sufficiently varied as to form an operator basis for the Hilbert space of the system, then the measurements are said to be *tomographically complete*, and furthermore, it becomes possible to derive an estimate for the identity of the quantum state from the empirical probabilities obtained by a process of *linear inversion* of the Born rule. One of the major drawbacks of the linear inversion method is that the resulting matrix need not be a valid quantum state; in particular, it may have negative eigenvalues and thus may assign, via the Born rule, negative probabilities to future measurement events. A particularly popular method which seeks to avoid the problem of negative eigenvalues is known as *maximum likelihood estimation*. The principle underlying maximum likelihood estimation is simply that the best estimate for the identity of an unknown quantum state should be one which maximizes the probability of the observed measurement data [[Hra97](#); [Hra+04](#)]. Although the method of maximum likelihood estimation always produces a positive semidefinite matrix as an estimate, it typically yields matrices which are rank-deficient in the sense that they assign *zero* probability to certain unobserved events; as Blume-Kohout

accurately argues, such a conclusion is theoretically unjustifiable after a finite number of trials [Blu10].

To avoid both the problems of negative and zero eigenvalues, one turns their attention to the *Bayesian* methods of quantum state estimation. The Bayesian approach to quantum state estimation seeks to determine an *a posteriori* belief about the identity of the quantum state based upon (i) statistical data obtained from a macroscopic measurement apparatus, and (ii) an *a priori* belief about the identity of the quantum state being measured [Hel69; Hol11; SBC01; Jon91; Buž+98]. In addition to the avoidance of zero and negative eigenvalues, the Bayesian approach to quantum tomography enables one to make statements about confidence regions [CR12]. There are at least two issues that emerge when incorporating prior knowledge about the identity of the quantum state in the context of quantum tomography. The first issue concerns the selection of a prior measure, while the second issue concerns the interpretation of the notion of an “unknown” quantum state.

Over the years, a number of principles have been developed for the purposes of determining a prior measure, from Laplace’s principle of *indifference* which seeks to identify priors which are considered, in some sense, uniform [Jef98] and later the principle of *invariant priors* which proposes the invariance of a prior under the action of a group of symmetries as a formalization of the notion of uniformity [Jon91; Har64; Jef98]. In the context of quantum theory, there are at least two cases to consider when selecting a prior measure over the space of states. If the state space is taken to be a homogeneous manifold upon which a compact group acts transitively, such as a finite-dimensional complex projective space equipped with a unitary group action, then the associated Haar measure, up to normalization, serves as the unambiguous invariant prior measure [Jon91; hai33]. On the other hand, when the state space is taken to include density operators, there exists no obvious symmetry group from which an invariant prior can be derived, and thus the treatment of the state space as a compact metric space permits an alternative notion of uniformity of the prior [Zyc+98; Bur69]. Alternatively, by appealing to the so-called purification postulate, one can propose priors over the space of density operators which are induced from priors over their purifications which are invariant with respect to the unambiguous Haar measure [Buž+98; TV99; ZS01].

Once a suitable prior has been chosen, there still remains an issue of interpretation; from the epistemological point of view that a quantum state is a description of an agent’s knowledge or belief about the outcomes of future measurement, the notion of taking many copies of an “unknown” quantum state becomes oxymoronic [CFS02]. Fortunately, this conceptual issue is satisfactorily resolved by quantum generalizations of de Finetti’s theorem from probability theory [FGP21]. While there are numerous de Finetti-type theorems in quantum theory [HM76; CFS02; KR05; Chr+07; Mit07; Chi10; LW17] their unifying characteristic is to build a formal bridge between (i) the operational notion of exchangeability or symmetry of an ensemble of states or measurements, and (ii) the algebraic notions of independence and convexity. More generally, quantum de Finetti theorems serve as the basis for a quantum theory of Bayesian inference [SBC01].

In practice, however, the number of measurements required to perform full quantum state tomography becomes unfeasible for large quantum systems [Aar07; Aar18; CW20]. In addition, in many contexts, one is merely interested in determining those properties of the quantum system which are functions of local, few-body observables [CW20; BBO20; ZRM21], or in certifying whether or not the unknown quantum state satisfies a particular

condition [MW13]. Consequently, a full reconstruction of the quantum state is often both unfeasible and unnecessary, and thus one seeks alternative measurement schemes which are optimized to produce only the information that is required [Bra+17; Aar18].

For example, suppose one is not interested in estimating the eigenvalues of an unknown quantum state, but merely its *spectrum* of eigenvalues. In 2001, Keyl and Werner demonstrated how the spectrum of a quantum state could be estimated from a projective measurement of a large number of copies of an unknown quantum state without knowing its corresponding eigenvectors [KW01]. Moreover, the authors demonstrated, for any given unknown state, the corresponding distribution of estimates satisfies the large deviations principle which quantifies the asymptotic rate of decay of the probabilities of incorrect estimates. In fact, Keyl and Werner’s paper on the topic of spectral estimation was perhaps the earliest and largest influence on the philosophical and technical ideas underlying this thesis. Beyond the obvious proposal of a projective measurement scheme for estimating the spectrum of a quantum state, Keyl and Werner’s result can also be understood as establishing a strong connection between the spectrum of a single quantum state and permutational symmetry on its many copies.<sup>1</sup>In recent years, this strong connection between spectra and permutational symmetry has been firmly established as a powerful theoretical tool. In particular, inequalities constraining the distribution of von Neumann entropies of a multipartite quantum state can be derived from corresponding constraints on the distribution of permutational symmetry [CM06; CŞW18].<sup>2</sup>

A few years later, Keyl generalized his large deviations approach to spectral estimation to the topic of full quantum state estimation [Key06; OW16]. Since then, these insights have been generalized further by Botero, Christandl, and Vrana [BCV21] and Franks and Walter [FW20] to consider the problem of estimating the *moment map* of an unknown quantum state. Loosely speaking, given a non-compact continuous Lie group,  $G$ , and a representation,  $(\Phi, \mathcal{H})$ , of that group acting on a Hilbert space  $\mathcal{H}$ , the *moment map* evaluated on a quantum state is a measure of how the *norm* of the state changes under the infinitesimal action of the group  $G$ . In particular, the moment map of a quantum state is simply its assignment of expectation values to the Hermitian observables in the Lie algebra of  $G$ . From this perspective, the problem of estimating the moment map of an unknown quantum state, with respect to a given representation, is a *generalization* of the problem of full quantum state estimation. In Refs. [BCV21; FW20], it was shown how to the moment map of an unknown quantum state, with respect to a fixed representation, could be estimated by performing a covariant measurement on  $n$  identical copies of that quantum state (in essentially the same spirit as Refs. [Chi10; Hol78; Mar12]). In particular, these measurement schemes emerge naturally from considering the  $n$ th tensor power representation along with its decomposition into its irreducible subrepresentations.

In general, the study of group actions on vector spaces, or more generally algebraic va-

---

<sup>1</sup>If the connection between spectra and permutational symmetry seems surprising, notice that the *purity*,  $\text{Tr}(\rho^2)$ , of a density matrix,  $\rho$ , (interpreted as a measure of concentration of a spectrum), is equivalent to a two-copy expectation value,  $\text{Tr}(X_{\text{swap}}\rho^{\otimes 2})$ , where  $X_{\text{swap}}$  is the operator which acts to permute the two copies of the underlying Hilbert space.

<sup>2</sup>This correspondence between quantum entropic inequalities on one hand and representation theoretic inequalities on the other can be seen as a quantum analogue of the seminal work of Chan and Yeung on the correspondence between Shannon inequalities and finite group inequalities [CY02] (see also [LC07]).

rieties, is the subject of geometric invariant theory [Woo10; Wal17; MN84; KN79; MFK94]. The connection between maximum likelihood estimation and concepts of stability from geometric invariant theory, has been previously developed by Améndola et al. for both Gaussian graphical models [Amé+21a] and discrete probabilistic models [Amé+21b]. Also note that techniques from geometric invariant theory have also been applied to related topics in quantum information theory including multipartite entanglement classification [Wal+13; Wer18; BRV18], canonical forms of tensor networks states [Acu+22], and quantum generalizations of the famous Brascamp-Lieb inequalities [BSW23; Gar+17; Ben+08].

## 1.2 Quantum realizability problems

A *quantum realizability*<sup>3</sup> *problem* refers to any decision problem which aims to determine whether or not there exists a quantum state which can simultaneously satisfy a given collection of constraints. Throughout this thesis, we have elected to conceptualize these constraints as describing potential properties a quantum state may or may not possess, and as such, a quantum realizability problem aims to characterize the relationships holding between the properties of quantum states. Furthermore, different examples of quantum realizability properties can be classified by considering the different collections of properties they pertain to. For the purposes of concreteness, next we endeavour to describe a small handful of motivating examples of quantum realizability problems.

**Uncertainty relations:** First and foremost, there exists a general class of quantum realizability problems which can be understood as a quantum generalization<sup>4</sup> of the multivariate moment problem which aims to characterize the relationships between the various statistical moments of multivariate probability distribution [KS13; Sto13]. For instance, as a special case, Heisenberg’s famous uncertainty relation holding between the variances associated to measurements of position and momentum observables [WZ14] can be understood as a necessary condition for the realizability for given values of variances for position and momentum. Similarly, Robertson’s uncertainty relation [Rob29], and later Schrödinger’s improvement [AB99] can be understood as necessary conditions for the realizability of a given collection of uncertainties and expectation values for a pair of observables and their commutators. Furthermore, if all of the properties under consideration are the variances (or equivalently uncertainties) associated to a given collection of observables, then the associated region of realizable uncertainties is known as the *uncertainty region* [Abb+16; BR19; Zha+22].

**Entanglement:** In the study of quantum entanglement, there are a few decision problems which may be interpreted as examples of quantum realizability problems. First, consider the problem of deciding whether or not a given bipartite quantum state is separable or entangled, which was shown to be an NP-hard problem by Gurvits [Gur04]. Although the

---

<sup>3</sup>Note the particular choice to use the adjectives “realizable” and “unrealizable” throughout this is merely our convention. Indeed, other authors have chosen alternative qualifying words, such as admissible/inadmissible, feasible/unfeasible, satisfiable/unsatisfiable, consistent/inconsistent, compatible/incompatible, or representable/unrepresentable.

<sup>4</sup>Here we are not referring to the seemingly related notion of a “quantum moment problem” as defined by [DPS04], but instead to realizability problems involving properties, and thus constraints, which are potentially polynomial functions of the underlying quantum state.



separability problem is not an example of a quantum realizability problem, it is related to an instance of a quantum realizability problem known as the symmetric extension problem [Che+14]. Given a positive integer  $k$  and bipartite quantum state  $\rho_{AB}$ , a  $(k+1)$ -partite quantum state,  $\sigma_{AB_1\dots B_k}$ , is said to be a  $k$ -symmetric extension of  $\rho_{AB}$  if it is (i) invariant under any permutation of  $k$  subsystems labelled by  $B$ , and (ii) satisfies  $\sigma_{AB_1} = \rho_{AB}$ . It can be shown that a bipartite quantum state is separable if and only if it admits of a symmetric extension for all positive integers  $k$  [DPS04]. From this perspective, any technique for verifying the non-existence of a symmetric extension can be used to verify the presence of bipartite entanglement.

A second example of a quantum realizability problem relating to entanglement is concerned with the existence of special quantum states which have the property of being *absolutely maximally entangled* [Hub17; Sco04; HC13]. A pure quantum state of  $n$ -qudits has the property of being  $m$ -uniform if all of its  $m$ -partite reduced states are maximally mixed. Furthermore, an  $m$ -uniform state is said to have the property of being absolutely maximally entangled whenever  $m = \lfloor n/2 \rfloor$ . For example, the two-qubit Bell-states are absolutely maximally entangled for  $n = 2$  and  $d = 2$ . The problem of deciding whether an absolutely maximally entangled state exists, for a given dimension  $d$  and number of qudits  $n$ , is thus an example of a quantum realizability problem. The existence of absolutely maximally entangled states is known to be directly related to the existence of quantum error correcting codes [Hub17; Sco04; Yu+21] as well as quantum secret-sharing schemes [HC13]. Unfortunately, despite recent progress concerning small dimensions and/or small numbers of qudits [HGS17; GW10], the existence of absolutely maximally entangled states, in full generality, remains an open problem.

**Distributed quantum entropies:** Another example of a quantum realizability problem is concerned with the allocation or distribution of von Neumann entropy throughout composite quantum systems [Pip03; LW05; Maj18]. Recall that the von Neumann entropy of a density operator, originally introduced in 1927 by von Neumann [Weh78; BŻ17], can be interpreted as a kind of quantum analogue to Gibb’s entropy from statistical mechanics or Shannon’s entropy from communication theory. Furthermore, Shannon’s entropy, using Shannon’s noiseless source coding theorem from communications theory, serves as a measure of the fundamental limit to data compressibility and thus as a measure of intrinsic information content [Sha48]. That there happens to be universal constraints on arrangement of entropies in composite systems, such as positivity, subadditivity, strong subadditivity and weak monotonicity [AL70], is well-known [Pip03; Maj18; Pet03]. The particular problem of deciding which allocations of von Neumann entropy are realizable by a quantum state is therefore an example of a quantum realizability problem wherein the realizable region, or rather its topological closure, is known to be a convex cone called the *entropy cone* [Pip03]. Nevertheless, despite being an active research question [Pip03; Kim20; Hay+04; Rus07] with many recent breakthroughs [LW05; Cad+14; CDW23], the joint realizability of a given collection of von Neumann entropies and their inequalities for  $n$ -partite quantum systems when  $n \geq 4$  remains an unresolved problem and a major open problem in quantum information theory.

**Quantum marginal problems:** The quantum marginal problem is the quantum analogue of a problem from probability theory, called the *classical* marginal problem. The classical marginal problem aims to characterize the relationships between the various marginal

distributions of a multivariate probability measure [FC12; Vor62; Mal88], and is intimately related to the derivation of entropic inequalities, obstructions in sheaf theory, and causal modelling [FC12; LSW11; AB11; FW18]. One of the earliest incarnations of the quantum marginal problem dates back to the late 1950s and early 1960s when, for the purposes of simplifying calculations of atomic and molecular structure, quantum chemists became interested in characterizing the possible reduced density matrices of a system of  $N$  interacting fermions [Cou60; Col63]. This version of the problem, known as the  $N$ -representability problem, has a long history [CY00; Col01; LTC+13; BD72; Rus07; Kly09] that continues to evolve [Maz12b; Maz12a; Kly06; Cas+21]. The quantum marginal problem aims to determine which collections of marginal quantum states, describing the configurations of differing quantum subsystems, can be understood as arising from some joint quantum state, describing the whole quantum system [TV15; Kla17; Hub17]. Variations of the quantum marginal problem arise when additional restrictions are placed on the form of the joint quantum state, e.g., by requiring the joint state to be fermionic [CY00; SGC13], bosonic [WMN10], Gaussian [Eis+08; Vla15], separable [NBA21], or having symmetric eigenvectors [AFT20]. In general, the quantum marginal problem has been shown to be a QMA-complete problem [Liu06; LCV07; WMN10; Boo12]. Using insights from representation theory and geometric invariant theory [BS00; Hec82], in the mid 2000s, Klyachko and Shumovsky completely solved the quantum marginal problem for disjoint subsystems [Kly04; KS06], which generalized earlier solutions for the case of a small number of low-dimensional subsystems [HSS03; Hig03; Bra03]. In particular, it was shown that the space of possible single-body quantum marginals, which depends only on the single-body spectra, forms a convex polytope, and thus is characterized by a finite set of linear inequality constraints. By comparison, when the quantum marginals pertain to overlapping subsystems, existing results are comparatively more sporadic and typically weaker, being only applicable to low-dimensional systems, small numbers of parties, or only yielding necessary but insufficient constraints [Che+14; CLL13; BSS06; Hal07; Che+16; CŞW18; DLN20]. Nevertheless, numerical methods for fully solving the general quantum marginal problem exist in the form of hierarchies of semidefinite programs [Yu+21], from which unrealizability witnesses can be extracted [Hal07].

**Methods:** Depending on the algebraic nature of the constraints under consideration, there are a number of different techniques which may be used to solve a given quantum realizability problem. For instance, semidefinite programming techniques can be readily be used to solve quantum realizability problems that pertain to quantum states described by a finite-dimensional density matrix subject to equality or inequality constraints which are *linear* functions of the candidate density operator [VB96]. Moreover, when the properties under consideration are *polynomial* functions of the underlying quantum state, it remains possible to construct a hierarchy of semidefinite programs which can approximately solve realizability problems which converge in some limit [BKM21; LGG21; LG22; NPA08]. Furthermore, when the properties under consideration are polynomial functions, techniques from computational algebraic geometry [CLO13] for performing non-linear quantifier elimination, such as cylindrical algebraic decomposition [Jir95], can, at least in principle, be used to analytically solve a given quantum realizability problem.

Alternatively, one can seek to characterize the relationships between properties of quantum states by probabilistic means; given a prior probability distribution over the space of quantum states, one can seek to derive the *induced* probability distribution over the space of



property values. This approach has been adopted for the purposes of characterizing the induced distribution of entanglement entropies of a bipartite pure state [Pag93], of expectation values of a single observable [VZ13], of expectation values of multiple observables [Zha+22; GZ13], of the reduced states of a bipartite pure state [ZS01], of eigenvalues of the one-body reduced states of a multipartite state [Chr+14] and of the marginals of a multipartite state [DLN20].

The purpose of this thesis is to explore an alternative method for tackling quantum realizability problems based upon the theory of property estimation outlined in Section 1.1. When the dimension of the Hilbert space is known and fixed, this method produces asymptotic conditions which are necessary and sufficient for the realizability of a given collection of properties of quantum states. Our primary application of this method is to the quantum marginal problem, where, in Chapter 6, we derive necessary and sufficient conditions for the realizability of any finite collection of candidate marginal quantum states. Although the evaluation of these conditions proves to be computationally challenging in general, it is our hope that, by building a conceptual bridge between property estimation theory and property realizability, future research will produce stronger and more tractable conditions.

## 1.3 Organization

The chapters of this thesis are largely intended to be read in chronological order as each chapter builds upon the insights gained from the previous chapter. The only two exceptions to this pattern are Chapter 2, which provides some mathematical background, and Chapter 6, which constitutes a standalone paper.

- **(Chapter 2) Preliminaries:** As this thesis relies heavily on the representation theory of finite-dimensional groups from the perspective of quantum theory and quantum measurements, we have elected to include a preliminary section to review the topics of measure theory, quantum theory, group theory and representation theory. Our presentation of representation theory focuses on the highest weight classifications of complex semisimple Lie algebras, compact connected Lie groups and their complexifications with the textbook by Hall as the main reference [Hal15]. Readers already familiar with these topics who wish to skip this preliminary chapter are encouraged to review a summary of our notational conventions in Section 2.5.
- **(Chapter 3) Non-commutative optimization:** This chapter is concerned with the geometric and algebraic aspects of the orbit of a vector under the action of group representation. Here we review the concepts of stability, capacity and norm minimization as well as the Kempf-Ness theorem which provides a deep connection between extremal surfaces of an orbit and the vanishing of the generalized gradient known as the moment map. The key result of this chapter is the strong duality theorem (Theorem 3.3.20) due to Franks and Walter [FW20], which, in later chapters, becomes the foundation for characterizing the asymptotic probabilities of quantum measurements applied to large ensembles of identical quantum states. The contents of this chapter are based partially on (i) the non-algorithmic aspects of the theory of non-commutative optimization due

to Bürgisser et al. [Bür+19], and (ii) the proof and interpretation of the strong duality result as a semiclassical limit due to Franks and Walter [FW20].

- **(Chapter 4) Estimation theory:** In this chapter we turn our attention to the topic of estimating various properties of quantum states. Given a fixed group representation, it is shown that its decomposition into irreducible subrepresentations naturally generates a covariant measurement which can be used to extract information about the covariant properties of the states they are applied to. In particular, by suitably deforming the strong duality result from Chapter 3, it is shown how the moment map of quantum state can be estimated by performing these representation-induced covariant measurements on a large number of independent copies of the state. As a special case, we recover the quantum state estimation result due to Keyl and Werner [KW01]. The contents of this chapter are based heavily on the works of Franks and Walter [FW20] and Botero, Christandl, and Vrana [BCV21].
- **(Chapter 5) Realizability:** The purpose of this chapter is to explore the connection between estimating the properties of a quantum state and determining the relationship between them. As such, this chapter relies on the moment map estimation result from Chapter 4, especially the deformed strong duality result encountered in Section 4.3. Our observation and guiding intuition is simple; a collection of candidate property values is *realizable* by some quantum state if and only if a random quantum state *occasionally* behaves as if it has those properties. From this principle, we recover the well-known result that the set of moment map values which can be realized by some quantum state forms a convex polytope known as the *moment polytope*. In addition, we apply this principle to obtain an asymptotic characterization of the jointly realizable region for a finite collection of moment maps, and demonstrate its potential applicability to a few open questions in quantum theory. The contents of this chapter represent partial progress toward generalizing the key ideas from Chapter 6, and are unpublished.
- **(Chapter 6) Quantum marginal problems:** This chapter is based entirely on the contents of my most recent paper, which makes partial progress on the aforementioned quantum marginal problem [Fra22]. As such, none of the previous chapters serve as prerequisites. Using simple principles of symmetry and operator positivity, we manage to derive a countable family of inequalities, each of which is necessarily satisfied by any realizable collection of quantum marginals. Additionally, we prove the sufficiency of this family of inequalities: every unrealizable collection of quantum marginals will violate at least one inequality belonging to the family.

## 1.4 A toy example

Here we present a toy example of a quantum realizability problem pertaining to a two-level quantum system, otherwise known as a qubit. Although this toy example admits of a rather straightforward solution, its purpose is to illustrate the variety of different approaches one might take in a more complicated scenario. Recall that a pure qubit quantum state,  $\psi$ , can

be faithfully described by,

$$|\psi\rangle = \cos\left(\frac{\theta}{2}\right)|0\rangle + e^{i\phi}\sin\left(\frac{\theta}{2}\right)|1\rangle, \quad (1.4.1)$$

for some  $\theta \in [0, \pi)$  and  $\phi \in [0, 2\pi)$ , or equivalently by a triple of coordinates,  $(x, y, z) \in \mathbb{R}^3$ , lying on the surface of the Bloch sphere subject to the constraint  $x^2 + y^2 + z^2 = 1$ . Moreover, the triple of coordinates  $(x, y, z)$  which describe the state  $\psi$  correspond precisely to the triple of expectation values  $(\langle X \rangle_\psi, \langle Y \rangle_\psi, \langle Z \rangle_\psi)$  where the observables  $X, Y$  and  $Z$  are the familiar Pauli matrices.

Now suppose, for the sake of exploring a toy problem, that one is interested in characterizing the relationship between just two of these observables, say the expectation values of  $X$  and  $Z$ . For the sake of notational convenience, consider three functions,  $e_X, e_Z$ , and  $e_{XZ}$ , of pure states  $\psi$  such that

$$e_X(\psi) = \langle X \rangle_\psi, \quad e_Z(\psi) = \langle Z \rangle_\psi, \quad e_{XZ}(\psi) = (\langle X \rangle_\psi, \langle Z \rangle_\psi). \quad (1.4.2)$$

When given a state,  $\psi$ , calculating the  $X$  and  $Z$  expectation values,  $e_{XZ}(\psi)$ , is a straightforward task. The inverse problem, however, is less straightforward; given a pair of values  $(x, z) \in \mathbb{R}^2$ , what is the corresponding pure state  $\psi$  such that  $(x, z) = e_{XZ}(\psi)$ ? What makes this inverse problem challenging is simply that sometimes no state exists, in which case  $(x, z) \in \mathbb{R}^2$  are said to be *unrealizable*. Additionally, even if the pair  $(x, z) \in \mathbb{R}^2$  is *realizable* as the  $X$  and  $Z$  expectation values of some state, the solution might not be unique.

**A geometric approach:** One strategy for describing the set of all possible  $X$  and  $Z$  expectation values, known as the *realizable region* is to appeal to the geometry of the Bloch sphere representation of qubit states. Since the coordinates,  $(x, y, z) \in \mathbb{R}^3$ , on the Bloch sphere correspond precisely to the triple of expectation values  $(\langle X \rangle, \langle Y \rangle, \langle Z \rangle)_\psi$ , one can readily conclude that the pair  $(x, z) \in \mathbb{R}^2$  is realizable as the pair of expectation values  $e_{XZ}(\psi) = (\langle X \rangle, \langle Z \rangle)_\psi$  of some pure qubit state  $\psi$  if and only if  $(x, z)$  lies inside the unit disk:

$$\exists \psi : e_{XZ}(\psi) = (\langle X \rangle, \langle Z \rangle)_\psi = (x, z) \iff x^2 + z^2 \leq 1. \quad (1.4.3)$$

Geometrically, the function  $e_{XZ}$  defined above can be viewed as an orthogonal projection of the Bloch sphere onto the  $(x, z)$ -plane in  $\mathbb{R}^2$  (see Fig. 1.1). The geometric approach also yields a characterization of the number of distinct solutions. Consider the set of all states  $\psi$  with  $X$  expectation value  $e_X(\psi) = x$ , denoted by  $e_X^{-1}(x)$ . Analogously, consider the sets  $e_Z^{-1}(x)$  and  $e_{XZ}^{-1}(x, z)$ . In particular, there is a useful relationship between these subsets of states:

$$e_{XZ}^{-1}(x, z) = e_X^{-1}(x) \cap e_Z^{-1}(z). \quad (1.4.4)$$

Geometrically, the sets  $e_X^{-1}(x)$  and  $e_Z^{-1}(z)$  are circular slices of the surface of the Bloch sphere and the number of intersections of these circles determines the number of states such that  $e_{XZ}(\psi) = (x, z)$  (see Fig. 1.2). The corresponding algebraic problem is to count the number of distinct real solutions to the equation  $x^2 + y^2 + z^2 = 1$  for fixed  $x$  and  $z$ . Since

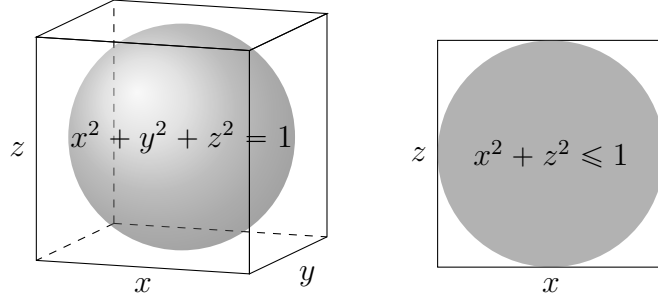


Figure 1.1: The realizable region for the  $\langle X \rangle_\psi$  and  $\langle Z \rangle_\psi$  expectation values of a pure qubit state arises from the orthogonal projection of the Bloch sphere onto the  $(x, z)$ -plane. In other words, there exists a pure state with expectation values  $(x, z) = (\langle X \rangle_\psi, \langle Z \rangle_\psi) \in \mathbb{R}^2$  if and only if  $x^2 + z^2 \leq 1$ .

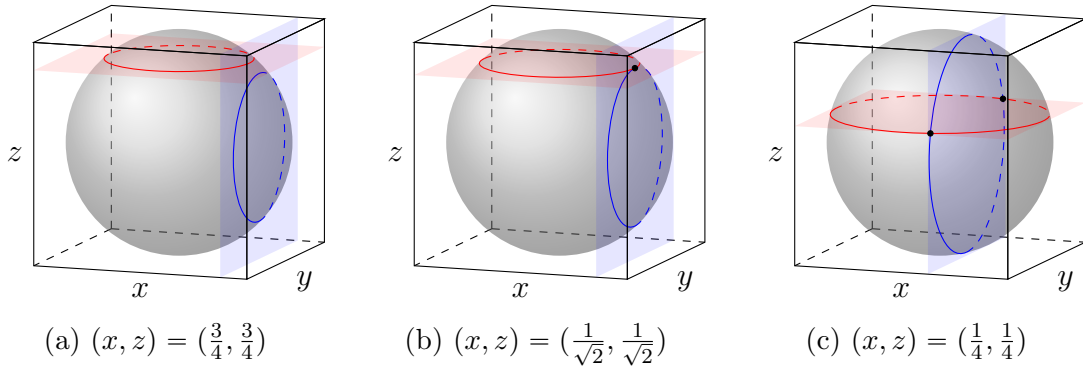


Figure 1.2: The intersection of  $e_X^{-1}(x)$  (blue) with  $e_Z^{-1}(z)$  (red) may have 0, 1 or 2 elements depending on the value of  $x^2 + z^2$ .

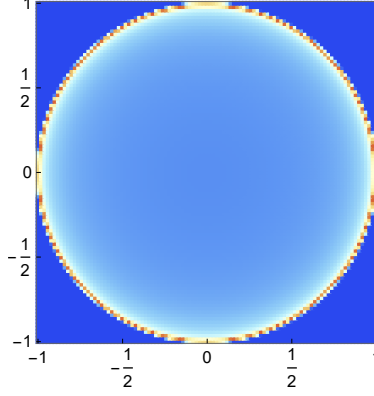


Figure 1.3: A visualization of pushforward measure  $\nu_{XZ} = \mu \circ e_{XZ}^{-1}$  for uniform prior  $\mu$ .

$y = \pm\sqrt{1 - x^2 - z^2}$ , the value of sign of  $1 - x^2 - z^2$  determines the cardinality of  $e_{XZ}^{-1}(x, z)$ :

$$|e_{XZ}^{-1}(x, z)| = \begin{cases} 0 & x^2 + z^2 > 1, \\ 1 & x^2 + z^2 = 1, \\ 2 & x^2 + z^2 < 1. \end{cases} \quad (1.4.5)$$

**A measure theory approach:** Another strategy for assessing the relationships between  $X$  and  $Z$  expectation values is to *measure* the volume of states whose expectation values belong to some region,  $e_{XZ}(\psi) \in \Delta \subset \mathbb{R}^2$ . Given a probability measure,  $\mu$ , over the set of states such as the uniform measure over the surface of the Bloch sphere,  $d\mu = (4\pi)^{-1} \sin \theta d\theta d\phi$ , the percentage of states,  $\psi$ , such that  $e_{XZ}(\psi) \in \Delta$  is given by the *pushforward measure*,  $\nu_{XZ} := \mu \circ e_{XZ}^{-1}$ , of  $\mu$  through  $e_{XZ}$ :

$$\nu_{XZ}(\Delta) := \mu(e_{XZ}^{-1}(\Delta)). \quad (1.4.6)$$

A direct calculation of the pushforward measure (Fig. 1.3) in this context reveals a probability density of the form,

$$d\nu_{XZ}(x, z) := \begin{cases} \frac{1}{2\pi\sqrt{1-x^2-z^2}} dx dz & x^2 + z^2 < 1, \\ 0 & x^2 + z^2 > 1. \end{cases} \quad (1.4.7)$$

Note that the support of the probability  $\nu_{XZ}$  is, perhaps unsurprisingly, equal to realizable region corresponding to the unit disk,  $x^2 + z^2 \leq 1$ . In other words, if  $\Delta \subset \mathbb{R}^2$  lies entirely outside of the unit disk, then  $\nu_{XZ}(\Delta) = 0$ .

**An estimation theory approach:** The estimation-theoretic approach naturally arises from acknowledging that expectation values are inherently statistical. To introduce the idea, note that an equivalent way to formalize the aforementioned pushforward measure is to reconceptualize the function  $e_{XZ}$ , which maps each state  $\psi$  to its pair of expectation values  $e_{XZ}(\psi) = (\langle X \rangle_\psi, \langle Z \rangle_\psi)$ , as a deterministic probability kernel, denoted by  $K_{XZ}$ , which maps

each state  $\psi$  to the point measure  $\delta_{e_{XZ}(\psi)}$ , concentrated at  $e_{XZ}(\psi)$ , such that for each region  $\Delta \in \mathbb{R}^2$ ,

$$K_{XZ}(\Delta|\psi) := \delta_{e_{XZ}(\psi)}(\Delta) = \begin{cases} 1 & e_{XZ}(\psi) \in \Delta, \\ 0 & e_{XZ}(\psi) \notin \Delta. \end{cases} \quad (1.4.8)$$

In this manner, the pushforward measure can be re-expressed as integration over  $K_{XZ}(\cdot|\psi)$  with respect to the prior measure  $\mu$ :

$$\nu_{XZ}(\Delta) = (\mu \circ e_{XZ}^{-1})(\Delta) = \int_{\psi} K_{XZ}(\Delta|\psi) d\mu(\psi). \quad (1.4.9)$$

The core idea of the estimation-theoretic approach is to *approximate* the pushforward measure  $\nu_{XZ}$  by approximating the probability kernel  $K_{XZ}(\Delta|\psi)$ , by performing a sufficiently large collective measurement on  $n$  copies of the state  $\psi$ , i.e.,

$$K_{XZ}(\Delta|\psi) \approx \langle \psi^{\otimes n}, E_n^{XZ}(\Delta) \psi^{\otimes n} \rangle, \quad (1.4.10)$$

where  $E_n^{XZ}$  is a quantum measurement, referred to as an *estimation scheme*, whose outcomes, when applied to  $\psi^{\otimes n}$ , correspond to estimates for the value of  $e_{XZ}(\psi)$ . By doing so, one obtains an approximation of the pushforward measure  $\nu_{XZ}$  of the form

$$\nu_{XZ}^{(n)}(\Delta) := \int_{\psi} d\mu(\psi) \langle \psi^{\otimes n}, E_n^{XZ}(\Delta) \psi^{\otimes n} \rangle. \quad (1.4.11)$$

Fortunately, examples of such sequences occur naturally in the context of estimating the expectation value of an unknown quantum state  $\psi$ . Perhaps the most natural way to jointly estimate both the  $X$  and  $Z$  expectation values of a state  $\psi$  is to partition the collection of prepared copies of  $\psi$  into two portions of roughly equal size and then, respectively on each portion, separately perform a projective measurement in the eigenbases of the observables  $X$  and  $Z$  and then let the corresponding empirical mean values serve as estimates for the value of  $e_{XZ}(\psi) = (\langle X \rangle_{\psi}, \langle Z \rangle_{\psi})$ . The resulting approximation for  $\nu_{XZ}(\Delta)$  is depicted in [Fig. 1.4](#).

**Realizability from occasionality:** Although the probability measure,  $\nu_{XZ}^{(n)}$ , based on an estimation-theoretic approach converges to the exact pushforward measure  $\nu_{XZ}^n = \mu \circ e_{XZ}^{-1}$  in the limit where  $n$  tends to infinity, for any finite  $n$ , there is always a chance to produce a pair of estimates  $(x, z)$  which are unrealizable. In other words, the positivity  $\nu_{XZ}^{(n)}(\{(x, z)\}) > 0$  is insufficient evidence to conclude that the pair  $(x, z)$  can be realized by some quantum state  $\psi$ . Nevertheless, the sequence of measures based on estimation theory,  $\nu_{XZ}^{(n)}$ , can still be used to distinguish between realizability and unrealizability.

To see how this might be possible, consider the scenario where  $n = 2m$  copies of a state  $\psi$  are prepared and the eigenbasis measurements of  $X$  and  $Z$  are respectively performed on  $m$  copies. Let the number of occurrences of spin-up, spin-down, spin-right and spin-left respectively be denoted by  $k_{\uparrow}, k_{\downarrow}, k_{\rightarrow}$  and  $k_{\leftarrow}$  such that  $k_{\uparrow} + k_{\downarrow} = k_{\rightarrow} + k_{\leftarrow} = m$ .

On the one hand, the likelihood of only observing spin-up or spin-right outcomes, and thus of producing the *unrealizable* estimate of  $e_{XZ}(\psi) \approx (+1, +1)$ , has the upper bound

$$\nu_{XZ}^{(2m)}(\{(+1, +1)\}) = \Pr(k_{\uparrow} = k_{\rightarrow} = m) \leq \left( \frac{3 + 2\sqrt{2}}{8} \right)^m. \quad (1.4.12)$$

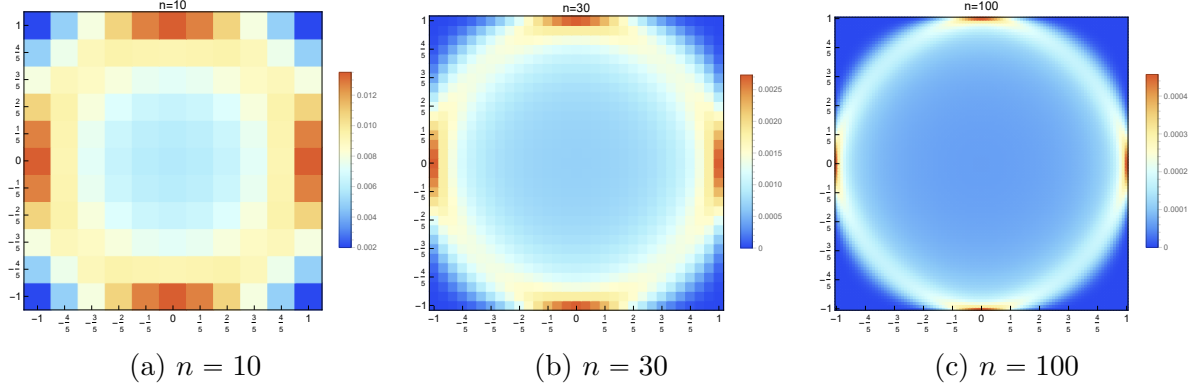


Figure 1.4: A sequence of approximations,  $\nu_{XZ}^{(2n)}$ , of the pushforward measure  $\nu_{XZ}$  in Fig. 1.3, obtained from performing the  $X$  eigenbasis measurement  $n$  times followed by the  $Z$  eigenbasis measurement  $n$  times on  $2n$  copies of a fixed state  $\psi$  sampled uniformly.

Therefore, the probability that a state  $\psi$  would behave in manner which would lead an experimenter to conclude  $\psi$  simultaneously satisfies both  $e_X(\psi) = \langle X \rangle_\psi = 1$  and  $e_Z(\psi) = \langle Z \rangle_\psi = 1$  decays to zero at an *exponential* rate with respect to the number of trials.

On the other hand, the likelihood of observing an equal number of spin-up and spin-down outcomes followed by an equal number of spin-right and spin-left outcomes, and thus of producing the *realizable* estimate of  $e_{XZ}(\psi) \approx (0, 0)$ , admits of lower-bound

$$\nu_{XZ}^{(2m)}(\{(0, 0)\}) = \Pr(k_\uparrow = k_\downarrow = k_\rightarrow = k_\leftarrow = \frac{m}{2}) \geq \begin{cases} 0 & m \text{ is odd,} \\ \frac{1}{2^m} & m \text{ is even.} \end{cases} \quad (1.4.13)$$

In contrast to the probability  $\nu_{XZ}^{(n)}(\{(+1, +1)\})$ , the probability  $\nu_{XZ}^{(2m)}(\{(0, 0)\})$  of producing the estimate  $(0, 0)$  does *not* decay to zero at an exponential rate with respect to the number of trials. Moreover, the lower-bound above can be interpreted as stating that the estimate  $(0, 0)$  is *occasionally* produced.

**A guiding principle:** Although the toy example of a quantum realizability problem presented in the section is somewhat contrived, its main purpose was to discover the following guiding principle which has served as the basis for tackling the less contrived examples:

---

*The sharp distinction between possibility and impossibility, or equivalently between realizability and unrealizability, can be faithfully captured by the comparatively fuzzy distinction between occasionality and exceptionality.*

# Chapter 2

## Preliminaries

This thesis aims to explore and develop a few ideas at the intersection of quantum theory, representation theory, and statistics. In order to support this exploration and development, this chapter endeavours to provide the unfamiliar reader with enough background to understand the results presented in subsequent chapters. For any reader already familiar with the aforementioned topics, the purpose of this chapter is to establish notational conventions, and moreover, to serve as a reminder of, and reference for, the following topics:

- **Measure theory:** measures & integration, probability theory & statistics, and types of convergence (Section 2.1).
- **Quantum theory:** the concept of a quantum state, observables, and quantum measurements (Section 2.2).
- **Group theory:** groups, subgroups, cosets, orbits, and group actions, with a focus on matrix Lie groups (Section 2.3).
- **Representation theory:** finite-dimensional representations, compactness & unitarity, reducibility & compositionality, roots & weights, complexification of Lie algebras, and the theorem of highest weights (Section 2.4).

This chapter concludes with a series of tables in Section 2.5 which summarize most of the commonly used notation throughout this thesis.

## 2.1 Measure theory

### 2.1.1 Measures

The purpose of this section is to provide a basic overview of all of the measure-theoretic terminology used throughout this thesis.

We begin with the usual definition of a  $\sigma$ -algebra [Tao11, Definition 1.4.12].

**Definition 2.1.1.** A  *$\sigma$ -algebra* on a set  $X$  is a non-empty collection  $\Sigma$  of subsets of  $X$  such that the following conditions hold.



- (i) (Empty set)  $\emptyset \in \Sigma$ .
- (ii) (Complements) If  $A \in \Sigma$ , then  $X \setminus A \in \Sigma$ .
- (iii) (Countable unions) If  $A_1, A_2, \dots \in \Sigma$ , then  $A_1 \cup A_2 \cup \dots \in \Sigma$ .

The pair  $(X, \Sigma)$  is referred to as a *measurable space*.

**Definition 2.1.2.** If  $X$  is a topological space, or rather  $(X, \tau)$  is a topological space with topology  $\tau$ , the *Borel  $\sigma$ -algebra*,  $\Sigma$ , is the smallest  $\sigma$ -algebra on  $X$  that contains  $\tau$ . The pair  $(X, \Sigma)$  is referred to as a *Borel space*.

A topological space  $(X, \tau)$  for which there exists a metric  $d : X \times X \rightarrow [0, \infty)$  which (i) induces the topology on  $X$ , and (ii) makes  $X$  both separable and complete as a metric space, is called a *Polish space*. The Borel space  $(X, \Sigma)$  associated to a Polish space is called a *standard Borel space*.

**Definition 2.1.3.** Let  $(X, \Sigma)$  be a measurable space. A function  $\mu : \Sigma \rightarrow [-\infty, \infty]$  is called a *non-negative measure* on  $X$  if it satisfies three conditions:

- *positivity*: for all  $\Delta \in \Sigma$ ,  $\mu(\Delta) \geq 0$ ,
- *nullity*:  $\mu(\emptyset) = 0$ , and
- *countable additivity*: for all countable collections  $\{\Delta_j \in \Sigma\}_{j \in \mathbb{N}}$ , of pairwise disjoint sets (meaning  $i \neq j \implies \Delta_i \cap \Delta_j = \emptyset$ ),  $\mu$  satisfies

$$\mu \left( \bigcup_{j \in \mathbb{N}} \Delta_j \right) = \sum_{j \in \mathbb{N}} \mu(\Delta_j). \quad (2.1.1)$$

If additionally,  $\mu(X) = 1$ , then  $\mu$  is called a *probability measure* on  $X$ .

**Definition 2.1.4.** Let  $(X, \Sigma(X))$  and  $(Y, \Sigma(Y))$  be measurable spaces. A function  $f : X \rightarrow Y$  is said to be *measurable* if for all  $A \in \Sigma(Y)$ ,  $f^{-1}(A) \in \Sigma(X)$ .

**Remark 2.1.5.** There are many equivalent ways of expressing the integration of a measurable real-valued function,  $f : X \rightarrow \mathbb{R}$ , with respect to a probability measure  $\mu : \Sigma(X) \rightarrow [0, 1]$ . For example, all of the following expressions are equivalent:

$$\mathbb{E}_\mu[f] = \int_X f \, d\mu = \int_{x \in X} f(x) \, d\mu(x) = \int_{x \in X} f(x) \mu(dx). \quad (2.1.2)$$

**Definition 2.1.6.** A *probability kernel* from  $(Y, \Sigma(Y))$  to  $(X, \Sigma(X))$  is a map  $\xi : \Sigma(X) \times Y \rightarrow [0, 1]$  such that

- i) for each  $y \in Y$ , the map  $\Delta \mapsto \xi(\Delta, y)$  is a probability measure on  $(X, \Sigma(X))$ , and
- ii) for each  $\Delta \in \Sigma(X)$ , the map  $y \mapsto \xi(\Delta, y)$  is  $\Sigma(Y)$ -measurable.

A probability kernel  $\xi : \Sigma(X) \times Y \rightarrow [0, 1]$  evaluated at  $(\Delta, y) \in \Sigma(X) \times Y$  may be alternatively written as  $\xi(\Delta|y)$  or  $\xi^y(\Delta)$ .

**Remark 2.1.7.** Associated to any measurable function  $f : \mathcal{S} \rightarrow X$  is a deterministic Markov kernel  $K_f : \Sigma(X) \times \mathcal{S} \rightarrow [0, 1]$  defined for  $\Delta \in \Sigma(X)$  as

$$K_f(\Delta, \rho) = \begin{cases} 1 & f(\rho) \in \Delta, \\ 0 & f(\rho) \notin \Delta. \end{cases} \quad (2.1.3)$$

**Definition 2.1.8** (The Pushforward of a Measure). Let  $f : \mathcal{S} \rightarrow X$  be a measurable function between measurable spaces  $(\mathcal{S}, \Sigma(\mathcal{S}))$  and  $(X, \Sigma(X))$ . For each probability measure  $\mu : \Sigma(\mathcal{S}) \rightarrow [0, 1]$ , there is an induced probability measure on  $X$  called the **pushforward measure** of  $\mu$  by  $f$ , denoted  $f_*\mu$ , and defined for  $\Delta \in \Sigma(X)$  as

$$(f_*\mu)(\Delta) := \mu(f^{-1}(\Delta)) = (\mu \circ f^{-1})(\Delta) = \int_{\rho \in \mathcal{S}} K_f(\Delta, \rho) \mu(d\rho). \quad (2.1.4)$$

## 2.1.2 Weak convergence

The purpose of this section is to define weak convergence of probability measures and then state, and sometimes prove, a small handful of related results that are used in the main text. Many of these definitions and results are taken directly from [DE11, Appendix A]. Throughout this section,  $X$  will denote a Polish space,  $\mathcal{P}(X)$  will be the set of probability measures on the standard Borel space  $(X, \Sigma(X))$ , and  $\mathcal{C}_b(X)$  will be the space of bounded, continuous functions from  $X$  to  $\mathbb{R}$ .

The following pair of definitions can be found in [DE11, Appendix A.3].

**Definition 2.1.9.** A sequence  $(\mu_n)_{n \in \mathbb{N}}$  in  $\mathcal{P}(X)$  **converges weakly** to  $\mu$  in  $\mathcal{P}(X)$ , denoted  $\mu_n \Longrightarrow \mu$ , if for all bounded continuous functions  $g \in \mathcal{C}_b(X)$ ,

$$\lim_{n \rightarrow \infty} \int_X g d\mu_n = \int_X g d\mu. \quad (2.1.5)$$

The notion of weak convergence can also be seen as a convergence with respect to a topology on  $\mathcal{P}(X)$ , namely the weak topology.

**Definition 2.1.10.** The **weak topology** is the topology on  $\mathcal{P}(X)$  generated by open neighborhoods around each  $\gamma \in \mathcal{P}(X)$  of the form

$$\left\{ \mu \in \mathcal{P}(X) \mid \left| \int_X g_i d\mu - \int_X g_i d\gamma \right| < \epsilon, i \in \{1, \dots, k\} \right\} \quad (2.1.6)$$

where  $\epsilon > 0$ ,  $k \in \mathbb{N}$  and  $g_i \in \mathcal{C}_b(X)$ .

There are a number of conditions that are equivalent to weak convergence; their equivalence is known as the Portmanteau Theorem [DE11, Thm. A.3.4].

**Theorem 2.1.11** (Portmanteau Theorem). *Let  $(\mu_n)_{n \in \mathbb{N}}$  be a sequence in  $\mathcal{P}(X)$  and let  $\mu \in \mathcal{P}(X)$ . The following are equivalent:*

- (i)  $\mu_n \Longrightarrow \mu$ .

(ii)  $\lim_{n \rightarrow \infty} \int_X g d\mu_n = \int_X g d\mu$  for all  $g$  bounded, uniformly continuous functions from  $X$  to  $\mathbb{R}$ .

(iii)  $\limsup_{n \rightarrow \infty} \mu_n(C) \leq \mu(C)$  for all closed  $C \in \Sigma(X)$ .

(iv)  $\liminf_{n \rightarrow \infty} \mu_n(O) \geq \mu(O)$  for all open  $O \in \Sigma(X)$ .

(v)  $\lim_{n \rightarrow \infty} \mu_n(A) = \mu(A)$  for  $A \in \Sigma(X)$  with  $\mu(\partial A) = 0$  where  $\partial A = \text{cl}A \setminus \text{int}A$  is the boundary of  $A$ .

**Lemma 2.1.12.** *Let  $h : X \rightarrow Y$  be a continuous map between Polish spaces and let  $h_* : \mathcal{P}(X) \rightarrow \mathcal{P}(Y)$  be the pushforward of probability measures, i.e., for  $\mu \in \mathcal{P}(X)$ , and  $\Delta \in \Sigma(Y)$ ,*

$$(h_*\mu)(\Delta) := \mu(h^{-1}(\Delta)). \quad (2.1.7)$$

Then  $h_*$  is continuous with respect to the weak topologies on  $\mathcal{P}(X)$  and  $\mathcal{P}(Y)$ .

*Proof.* Let  $g \in \mathcal{C}_b(Y)$ , and consider a sequence  $(\mu_n)_{n \in \mathbb{N}}$  in  $\mathcal{P}(X)$  converging weakly to  $\mu$ . Then by the continuity of  $h$ ,  $g \circ h \in \mathcal{C}_b(X)$  and thus  $(h_*\mu_n)_{n \in \mathbb{N}}$  converges weakly to  $h_*\mu$ :

$$\int_Y g d(h_*\mu_n) = \int_X (g \circ h) d\mu_n \implies \int_X (g \circ h) d\mu = \int_Y g d(h_*\mu). \quad (2.1.8)$$

□

The next series of results are concerned with sequences of probability measures  $(\xi_n^y)_{n \in \mathbb{N}}$  indexed by  $y \in Y$  where  $Y$  is some other Polish space. Under suitable conditions, point-wise weak convergence of  $\xi_n^y$  for each  $y \in Y$  implies weak convergence of  $(\int_Y d\mu \xi_n^y)_{n \in \mathbb{N}}$  for some probability measure  $\mu \in \mathcal{P}(Y)$ . Of course, in order the integration to make sense, the maps  $y \mapsto \xi_n^y(\Delta)$  must be measurable.

The following lemma appears as [DE11, Theorem A.5.8].

**Lemma 2.1.13.** *Let  $\mu : \Sigma(Y) \rightarrow [0, 1]$  be a probability measure, let  $(\xi_n : \Sigma(X) \times Y \rightarrow [0, 1])_{n \in \mathbb{N}}$  be a sequence of probability kernels and  $\xi : \Sigma(X) \times Y \rightarrow [0, 1]$  a probability kernel. Assume that, for each  $y \in Y$ , the sequence  $(\xi_n^y)_{n \in \mathbb{N}}$  converges weakly to  $\xi^y$ . Then the sequence of measures  $\xi_n(dx|y) \otimes \mu(dy)$  converges weakly to  $\xi(dx|y) \otimes \mu(dy)$ .*

$$\xi_n(dx|y) \otimes \mu(dy) \implies \xi(dx|y) \otimes \mu(dy). \quad (2.1.9)$$

*Proof.* Since  $Y$  and  $X$  are assumed to be Polish spaces, to prove the weak convergence claimed above, by [DE11, Theorem A.3.14], it is sufficient to prove that the limit in Definition 2.1.9 holds for bounded, continuous functions  $g \in \mathcal{C}_b(X \times Y)$  of the form  $g(x, y) = s(x)t(y)$  for  $s \in \mathcal{C}_b(X)$  and  $t \in \mathcal{C}_b(Y)$ .

$$\lim_{n \rightarrow \infty} \int_{X \times Y} s(x)t(y) \xi_n(dx|y) \otimes \mu(dy) = \lim_{n \rightarrow \infty} \int_Y \left( \int_X s(x) \xi_n(dx|y) \right) t(y) \mu(dy) \quad (2.1.10)$$

Define  $q_n(y)$  for  $y \in Y$  as

$$q_n(y) := \int_X s(x) \xi_n(dx|y), \quad (2.1.11)$$

and note that  $q_n(y)$  is bounded point-wise for each  $y$  independently of  $n$  (since  $s(x)$  is bounded and  $\xi_n$  is a probability kernel). Therefore, the Lebesgue's dominated convergence theorem applies, and therefore

$$\lim_{n \rightarrow \infty} \int_{X \times Y} s(x)t(y)\xi_n(dx|y) \otimes \mu(dy) \quad (2.1.12)$$

$$= \int_Y \lim_{n \rightarrow \infty} \left( \int_X s(x)\xi_n(dx|y) \right) t(y)\mu(dy) \quad (2.1.13)$$

$$= \int_Y \left( \int_X s(x)\xi(dx|y) \right) t(y)\mu(dy) \quad (2.1.14)$$

$$= \int_{X \times Y} s(x)t(y)\xi(dx|y) \otimes \mu(dy). \quad (2.1.15)$$

Thus the theorem holds.  $\square$

**Corollary 2.1.14.** *Let everything be defined as in Lemma 2.1.13. For each  $n \in \mathbb{N}$ , define the probability measure  $\Xi_n : \Sigma(X) \rightarrow [0, 1]$  for each  $\Delta \in \Sigma(X)$  as*

$$\Xi_n(\Delta) := \int_Y \xi_n^y(\Delta)\mu(dy), \quad (2.1.16)$$

and similarly for  $\Xi : \Sigma(X) \rightarrow [0, 1]$ . Then  $(\Xi_n)_{n \in \mathbb{N}}$  converges weakly to  $\Xi$ , i.e.  $\Xi_n \Longrightarrow \Xi$ .

**Corollary 2.1.15.** *Let everything be defined as in Lemma 2.1.13 and Corollary 2.1.14. Assume there exists a measurable function  $f : Y \rightarrow X$  such that, for each  $y \in Y$ ,  $(\xi_n^y)_{n \in \mathbb{N}}$  converges weakly to  $\delta_{f(y)}$ , i.e.  $\xi_n^y \Longrightarrow \delta_{f(y)}$ . Then  $(\Xi_n)_{n \in \mathbb{N}}$  converges weakly to the pushforward measure  $f_*\mu$ , i.e.*

$$\Xi_n \Longrightarrow f_*\mu \quad (2.1.17)$$

*Proof.* First note that  $\delta_{f(\cdot)}(\cdot) : \Sigma(X) \times Y \rightarrow [0, \infty]$  is indeed a probability kernel because  $y \mapsto \delta_{f(y)}(\Delta)$  is  $\Sigma(Y)$ -measurable for each fixed  $\Delta \in \Sigma(X)$ . Then the corollary follows from an application of Corollary 2.1.14:

$$\Xi_n(dx) = \int_Y \xi_n(dx|y)\mu(dy) \Longrightarrow \int_Y \delta_{f(y)}(dx)\mu(dy) = \mu(f^{-1}(dx)) = (f_*\mu)(dx). \quad (2.1.18)$$

$\square$

We now turn our attention to the evaluation of a sequence of probability measures  $(\mu_n : \Sigma(X) \rightarrow [0, 1])_{n \in \mathbb{N}}$  along sequences of Borel sets  $(\Delta_n)_{n \in \mathbb{N}}$  that, in some sense, converge, or zoom into, to a single point  $x \in X$ .

**Definition 2.1.16.** Let  $X$  be a metric space. A sequence  $(\Delta_n \subseteq X)_{n \in \mathbb{N}}$  of subsets in  $X$  is said to be **nested** if

$$\forall n \in \mathbb{N} : \Delta_{n+1} \subseteq \Delta_n. \quad (2.1.19)$$

A **Cantor sequence** is a nested sequence of non-empty, closed subsets which satisfy

$$\lim_{n \rightarrow \infty} \text{diam}(\Delta_n) = 0, \quad (2.1.20)$$

where  $\text{diam}(S)$  is the **diameter** of  $S \subseteq X$ :

$$\text{diam}(S) = \sup\{d(x, x') \mid x, x' \in S\}. \quad (2.1.21)$$

We refer to such sequences as Cantor sequences because of the following well-known theorem, known as Cantor's intersection theorem for complete metric spaces

**Theorem 2.1.17** (Cantor's Intersection Theorem). *Let  $(\Delta_n \subseteq X)_{n \in \mathbb{N}}$  be a Cantor sequence (see [Definition 2.1.16](#)) in complete metric space  $X$ . Then the intersection over all  $\Delta_n$  is non-empty and contains exactly one point:*

$$\bigcap_{n \in \mathbb{N}} \Delta_n = \{x\} \quad (2.1.22)$$

for some  $x \in X$ .

*Proof.* See [[Rud53](#), Chap. 2]. □

In light of [Theorem 2.1.17](#), a Cantor sequence  $(\Delta_n)_{n \in \mathbb{N}}$  such that  $\bigcap_{n \in \mathbb{N}} \Delta_n = \{x\}$  will also be referred to a Cantor sequence converging to  $x$ . Also note that a Cantor sequence  $(\Delta_n)_{n \in \mathbb{N}}$  converging to  $x$  also forms a neighborhood basis for  $x$ , in the sense that every neighborhood  $N$  of  $x$  eventually contains  $\Delta_n$  for sufficiently large  $n$ .

**Lemma 2.1.18** (Typicality). *Let  $(\mu_n)_{n \in \mathbb{N}}$  be a sequence of probability measures converging weakly to the Dirac measure,  $\delta_x$ , concentrated at  $x \in X$ . Then there exists a Cantor sequence  $(\Delta_n)_{n \in \mathbb{N}}$  converging to  $x$  such that*

$$\lim_{n \rightarrow \infty} \mu_n(\Delta_n) = 1. \quad (2.1.23)$$

*Proof.* For  $\epsilon > 0$ , define

$$B_\epsilon(x) = \{x' \in X \mid d(x', x) < \epsilon\}. \quad (2.1.24)$$

Since  $(\mu_n)_{n \in \mathbb{N}}$  converges weakly to the point measure  $\delta_x$  at  $x$ , by condition (iii) of [Theorem 2.1.11](#), there exists a finite  $N_\epsilon \in \mathbb{N}$  such that,

$$\forall n \geq N_\epsilon : \mu_n(B_\epsilon(x)) \geq \delta_x(B_\epsilon(x)) - \epsilon = 1 - \epsilon. \quad (2.1.25)$$

Without loss of generality, we can assume that  $N_\epsilon$  is the minimum  $n \in \mathbb{N}$  such that [Eq. \(2.1.25\)](#) holds. Note that  $\epsilon \geq \epsilon'$  implies  $B_\epsilon(x) \supseteq B_{\epsilon'}(x)$  and thus  $N_\epsilon \leq N_{\epsilon'}$ . Next define a sequence  $q(n)$  by

$$q(n) = \min\{m^{-1} \mid m \in \mathbb{N}, n \geq N_{m^{-1}}\}. \quad (2.1.26)$$

If for any finite  $n' \in \mathbb{N}$ ,  $q(n')$  does not exist (i.e., the minimum above does not exist), then one can already conclude that the lemma holds because  $\mu_n(\{x\}) = 1$  for all  $n \geq n'$  (because measures on Polish spaces are necessarily inner regular). Therefore, we may proceed with the case where the minimum above always exists.

For each  $n$ , let  $\Delta_n$  be the closure of the smallest open ball around  $x$  with radius  $\frac{1}{m}$  for some  $m \in \mathbb{N}$  that satisfies, for all  $n \in \mathbb{N}$ , the equation

$$\mu_n(\Delta_n) \geq \mu_n(B_x(q(n))) = \mu_n(B_x(m^{-1})) \geq 1 - \frac{1}{m} = 1 - q(n). \quad (2.1.27)$$

Critically, since  $q(n)$  is non-increasing with respect to increasing  $n$ , we have  $\Delta_{n+1} \subseteq \Delta_n$  and furthermore  $\lim_{n \rightarrow \infty} q(n) = 0$ . As a consequence, [Theorem 2.1.17](#) holds and thus  $\bigcap_{n \in \mathbb{N}} \Delta_n = \{x\}$  and

$$\lim_{n \rightarrow \infty} \mu_n(\Delta_n) = 1 \quad (2.1.28)$$

hold simultaneously. □

### 2.1.3 Moments & cumulants

**Remark 2.1.19.** Let  $R$  be an  $\mathbb{R}$ -valued random variable. The following functions of  $t \in \mathbb{R}$  encode statistical properties of  $R$ . These are

- i) the *moment* generating function  $M(t) := \mathbb{E}[\exp(tR)]$ ,
- ii) the (*first*) *cumulant* generating function  $K(t) := \log \mathbb{E}[\exp(tR)]$ ,
- iii) the *characteristic* function  $\varphi(t) := \mathbb{E}[\exp(itR)]$ ,
- iv) and the (*second*) *cumulant* generating function  $H(t) := \log \mathbb{E}[\exp(itR)]$ .

Evidently these expressions are related by formal substitutions of the form  $t \mapsto \pm it$  or alternatively  $R \mapsto \pm iR$ .

**Definition 2.1.20.** Let  $M(t)$  be the moment generating function of a real-valued random variable  $R$ :

$$M(t) = \mathbb{E}[\exp(tR)]. \quad (2.1.29)$$

The  *$n$ th moment* of the random variable  $R$ , denoted  $\mu_n$ , is the derivative of  $M$  evaluated at  $t = 0$ :

$$\mu_n := \mathbb{E}[R^n] = (\partial_t^n M)(0). \quad (2.1.30)$$

The cumulant generating function  $K$  is simply the logarithm of the moment generating function. The  *$n$ th cumulant* of the random variable  $R$ , denoted  $\kappa_n$ , is analogously the derivative of  $K$  evaluated at  $t = 0$ :

$$\kappa_n := (\partial_t^n K)(0). \quad (2.1.31)$$

**Remark 2.1.21.** The first three derivatives of the cumulant generating function  $K$  in relation to the derivatives of the moment generating function are

$$\begin{aligned} \partial_t K &= \frac{\partial_t M}{M}, & \partial_t^2 K &= \frac{M \partial_t^2 M - (\partial_t M)^2}{M^2}, \\ \partial_t^3 K &= \frac{M^2 \partial_t^3 M - 3M(\partial_t^2 M)(\partial_t M) + 2(\partial_t M)^3}{M^3}. \end{aligned} \quad (2.1.32)$$

Evaluating everything at  $t = 0$  (where  $\mu_0 = 1$  and  $\kappa_0 = 0$ ) reveals

$$\kappa_1 = \mu_1 = \mathbb{E}[R], \quad \kappa_2 = \mu_2 - \mu_1^2 = \mathbb{E}[R^2] - \mathbb{E}[R]^2, \quad (2.1.33)$$

$$\kappa_3 = \mu_3 - 3\mu_2\mu_1 + 2\mu_1^3 = \mathbb{E}[R^3] - 3\mathbb{E}[R^2]\mathbb{E}[R] + 2\mathbb{E}[R]^3, \quad (2.1.34)$$

$$\begin{aligned} \kappa_4 &= \mu_4 - 3\mu_2^2 - 4\mu_1\mu_3 + 12\mu_1^2\mu_2 - 6\mu_1^4, \\ &= \mathbb{E}[R^4] - 3\mathbb{E}[R^2]^2 - 4\mathbb{E}[R]\mathbb{E}[R^3] + 12\mathbb{E}[R]^2\mathbb{E}[R^2] - 6\mathbb{E}[R]^4. \end{aligned} \quad (2.1.35)$$

## 2.1.4 Rate functions

The subject of large deviation theory aims to describe the asymptotic behaviour of sequences of probability measures,  $(\mu_n : \Sigma(X) \rightarrow [0, 1])_{n \in \mathbb{N}}$  over a shared measurable space  $(X, \Sigma(X))$  which, for sufficiently large  $n$ , assigns vanishingly small probability to certain events which might be considered as a large deviation from the expected event. In particular, the probabilities associated to these large deviations decay to zero as  $n$  tends to infinity at an *exponential* rate. As the magnitude of this rate of decay depends on the degree of deviation, large deviation theory introduces the concept of a rate *function* which assigns, to each  $x \in X$ , the rate  $I(x) \in [0, \infty]$ .

Unless otherwise stated, the sample space  $X$  will always be considered to be a Polish space (recall [Definition 2.1.2](#)) and thus the associated Borel measurable space,  $(X, \Sigma(X))$ , is a *standard Borel space*.

**Definition 2.1.22** (Rate Functions). A function  $I : X \rightarrow [0, \infty]$  is called a *rate function* if for all values of  $c \in [0, \infty)$ , the *lower level set*,  $L_I(c)$ , defined by

$$L_I(c) := \{x : I(x) \leq c\} \subseteq X, \quad (2.1.36)$$

is a compact subset of  $X$ .

Note that some references only require that the lower level sets of a rate function be closed, or equivalently (by [Lemma A.1.1](#)), that a rate function is lower-semicontinuous. In these references, a rate function with compact lower level sets is given the adjective *good* [[DZ10](#)] or less commonly *regular* [[LS87](#)]. In this paper, all rate functions will be shown to have compact lower level sets for the ease of exposition, thus following the convention of [[DE11](#)]. In cases where the space  $X$  is itself compact (which is the case for many property spaces of finite-dimensional quantum states), every closed subset is also compact and thus the distinction becomes unnecessary to make.

In general, the compactness of the lower-level sets of a rate function ensures that the infimum of a rate function,  $\inf_{x \in \Delta} I(x)$ , when taken over closed subsets  $\Delta \subseteq X$ , is always attained by some element  $x \in \Delta$ .

## 2.1.5 Principles of large deviations

Just as there are multiple equivalent ways to formulate the notion of weak convergence of sequences probability measures using the Portmanteau theorem ([Theorem 2.1.11](#)), as noted by Dupuis and Ellis, there are multiple ways to formulate the central notions of large deviation theory [[DE11](#); [LS87](#); [DZ10](#); [Den08](#)]. For a physicist-friendly introduction to the subject of large deviation theory and the foundational role it plays in statistical mechanics, see [[Tou11](#); [Tou11](#)].

The first notion of large deviation theory is the idea of a sequence of probability measures being tightly concentrated over compact subsets.

**Definition 2.1.23.** Let  $(X, \Sigma(X))$  be a standard Borel space. A sequence of probability measures  $(\mu_n : \Sigma(X) \rightarrow [0, 1])_{n \in \mathbb{N}}$  is said to be *exponentially tight* if for all  $r < \infty$  there exists a compact subset  $K_r \in \Sigma(X)$  such that

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \mu_n(X \setminus K_r) < -r. \quad (2.1.37)$$

An exponentially tight sequence of probability measures has the property that for any rate  $r > 0$ , you can find a compact subset  $K_r \in \Sigma(X)$  and a sufficiently large  $n \in \mathbb{N}$  such that

$$\mu_n(K_r) > 1 - \exp(-nr/2). \quad (2.1.38)$$

A stronger requirement<sup>1</sup> than exponential tightness is the notion of the large deviation principle [DZ10].

**Definition 2.1.24** (Large Deviation Principle). Let  $(X, \Sigma(X))$  be a standard Borel space. A sequence of probability measures  $(\mu_n : \Sigma(X) \rightarrow [0, 1])_{n \in \mathbb{N}}$  satisfies the **large deviation principle (LDP)** with rate function  $I : X \rightarrow [0, \infty]$  if it satisfies

(i) the **LDP upper bound**: for each closed subset  $C \in \Sigma(X)$ ,

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log \mu_n(C) \leq - \inf_{x \in C} I(x), \quad \text{and} \quad (2.1.39)$$

(ii) the **LDP lower bound**: for each open subset  $O \in \Sigma(X)$ ,

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \log \mu_n(O) \geq - \inf_{x \in O} I(x). \quad (2.1.40)$$

Before continuing, it is worth noticing that if a sequence of probability measures  $(\mu_n : \Sigma(X) \rightarrow [0, 1])_{n \in \mathbb{N}}$  satisfies the Large deviation principle upper bound for some rate function  $I : X \rightarrow [0, \infty]$ , then letting  $C = X$ , we conclude that the minimum of  $I$  over all  $X$  is exactly zero as  $\mu_n$  (because  $\mu_n$  is normalized and thus  $\log \mu_n(X) = 0$ ). Our first result is concerned with the special case where the rate functions vanish at a single point in  $X$ .

**Lemma 2.1.25.** *Let  $(X, \Sigma(X))$  be a standard Borel space and let  $(\mu_n : \Sigma(X) \rightarrow [0, 1])_{n \in \mathbb{N}}$  be a sequence of probability measures satisfying the large deviation principle upper bound with rate function  $I : X \rightarrow [0, \infty]$ . Assume that the rate function vanishes at a single point  $x \in X$ . Then the sequence  $(\mu_n)_{n \in \mathbb{N}}$  converges weakly to the Dirac measure,  $\delta_x$ , concentrated at  $x$ .*

*Proof.* Note that by the Portmanteau theorem [Theorem 2.1.11](#),  $\mu_n \implies \delta_x$  is equivalent to the claim that for all closed subsets  $C \subseteq X$ ,

$$\limsup_{n \rightarrow \infty} \mu_n(C) \leq \delta_x(C). \quad (2.1.41)$$

If  $x \in C$ , then  $\delta_x(C) = 1$  and thus the above bound would hold simply because  $\mu_n$  is a probability measure. What remains to prove, therefore, is the case where  $x \notin C$ . When  $x \notin C$ ,  $\delta_x(C) = 0$  and thus the limit above becomes equivalent to the following condition:

$$\lim_{n \rightarrow \infty} \mu_n(C) = 0. \quad (2.1.42)$$

---

<sup>1</sup>That satisfaction of the large deviation principle implies exponential tightness follows from our assumptions that (i)  $X$  is a Polish space and (ii) rate functions have compact lower-level sets [DZ10, Pg. 8].



To prove this limit holds, note that the sequence  $(\mu_n)_{n \in \mathbb{N}}$  satisfies the large deviation principle upper bound with rate function  $I : X \rightarrow [0, \infty]$ . Therefore, for all closed subsets  $C \subseteq X$ ,

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log \mu_n(C) \leq -I(C), \quad \text{where} \quad I(C) := \inf_{x \in C} I(x). \quad (2.1.43)$$

Consequently, for any  $\epsilon > 0$ , there exists an  $N \in \mathbb{N}$  such that for all  $n \geq N$ ,

$$\frac{1}{n} \log \mu_n(C) \leq -I(C) + \epsilon. \quad (2.1.44)$$

Since  $C$  is closed, and  $I$  has compact lower level sets, by the extreme value theorem,  $I(C) = \inf_{x \in C} I(x)$  is attained by some  $x^* \in C$ . Since  $I$  only vanishes at  $x$  and  $x \notin C$ , we can conclude that  $I(C) > 0$  is strictly positive. Taking  $\epsilon = I(C)/2 > 0$  yields for all  $n \geq N$ ,

$$\mu_n(C) \leq \exp(-nI(C)/2). \quad (2.1.45)$$

Therefore, taking the limit as  $n \rightarrow \infty$  yields Eq. (2.1.42) and therefore the lemma holds.  $\square$

Once it has been established that a sequence of measures  $(\mu_n : \Sigma(X) \rightarrow [0, 1])_{n \in \mathbb{N}}$  satisfies the large deviation principle with rate function  $I : X \rightarrow [0, \infty]$ , the contraction principle enables one to prove the large deviation principle holds sequences obtained through the pushforward of a continuous function [DZ10, Theorem 4.2.1].

**Proposition 2.1.26** (Contraction principle). *Let  $X$  and  $Y$  be Hausdorff topological spaces and  $g : X \rightarrow Y$  a continuous function. Consider a rate function  $I : X \rightarrow [0, \infty]$  and define  $I' : Y \rightarrow [0, \infty]$  for each  $y \in Y$  with  $g^{-1}(y)$  non-empty by*

$$I'(y) = \inf\{I(x) \mid x \in X, y = g(x)\}. \quad (2.1.46)$$

*and otherwise  $I'(y) = \infty$ . Then  $I'$  is a rate function. Moreover, if  $(\mu_n)_{n \in \mathbb{N}}$  satisfies the LDP with rate function  $I$ , then  $(\mu_n \circ g^{-1})_{n \in \mathbb{N}}$  satisfies the LDP with rate function  $I'$ .*

## 2.2 Quantum theory

### 2.2.1 Quantum states

While the state of a system in quantum theory can, depending on the particular application, be modeled by a variety of different mathematical objects, the notion of a *Hilbert space*, which is a complete inner product space, plays a central role. Throughout this thesis, the only kind of Hilbert space that will be considered will be a complex, finite-dimensional Hilbert space.

**Definition 2.2.1.** A **complex finite-dimensional Hilbert space**, denoted  $\mathcal{H}$ , is a complex vector space,  $\mathcal{H}$ , of dimension  $d = \dim(\mathcal{H}) < \infty$ , equipped with a sesquilinear inner product  $\langle \cdot, \cdot \rangle : \mathcal{H} \times \mathcal{H} \rightarrow \mathbb{C}$  which induces the norm  $\|\cdot\| : \mathcal{H} \rightarrow \mathbb{R}_{\geq 0}$  defined by

$$\|v\| := \sqrt{\langle v, v \rangle}. \quad (2.2.1)$$

**Example 2.2.2.** The canonical example of a complex  $d$ -dimensional Hilbert space is the set  $\mathbb{C}^d$  of all tuples  $(z_1, \dots, z_d)$  of  $d$  complex numbers, together with the standard basis  $\{e_1, \dots, e_d\}$  where  $e_j$  is the tuple of all zeros except for a one in the  $j$ th position,

$$e_j := (0, \dots, \overset{j-1}{\underset{\cdot}{0}}, 1, 0, \dots, \overset{d-j}{\underset{\cdot}{0}}, 0). \quad (2.2.2)$$

The standard inner product,  $(\cdot, \cdot)$ , is defined by

$$(v, w) = \sum_{j=1}^d v_j^* w_j \quad (2.2.3)$$

where  $z^*$  denotes the complex conjugation of the complex number  $z$ .

**Definition 2.2.3.** A basis  $\{e_1, \dots, e_d\} \subset \mathcal{H}$  of a  $d$ -dimensional complex Hilbert space is **orthogonal** if

$$\forall i, j : \langle e_i, e_j \rangle = 0, \quad (2.2.4)$$

and **orthonormal** if, in addition to being orthogonal, it satisfies for all  $i$ ,  $\|e_i\| = \sqrt{\langle e_i, e_i \rangle} = 1$ .

**Definition 2.2.4.** Given a complex finite-dimensional Hilbert space,  $\mathcal{H}$ , the set of linear mappings from  $\mathcal{H}$  to itself will be denoted by  $\text{End}(\mathcal{H})$ . The **\*-involution** or **adjoint** on  $\text{End}(\mathcal{H})$  is a linear map sending  $X \in \text{End}(\mathcal{H})$  to the unique linear map  $X^* \in \text{End}(\mathcal{H})$  satisfying for all  $v, w \in \mathcal{H}$  the equation

$$\langle X^* v, w \rangle = \langle v, X w \rangle. \quad (2.2.5)$$

This map is an involution in the sense that  $(X^*)^* = X$  for all  $X \in \text{End}(\mathcal{H})$ .

**Example 2.2.5.** Given an orthonormal basis  $\{e_1, \dots, e_d\}$  for  $\mathcal{H}$ , define the **trace** operation as the linear map

$$\text{Tr} : \text{End}(\mathcal{H}) \rightarrow \mathbb{C}, \quad (2.2.6)$$

defined for all  $X \in \text{End}(\mathcal{H})$  by

$$\text{Tr}(X) := \sum_{j=1}^d \langle e_j, X e_j \rangle. \quad (2.2.7)$$

The **Hilbert-Schmidt inner product** is an inner product on  $\text{End}(\mathcal{H})$  defined by

$$\langle X, Y \rangle_{\text{HS}} := \text{Tr}(X^* Y) = \sum_{j=1}^d \langle e_j, X^* Y e_j \rangle = \sum_{j=1}^d \langle X e_j, Y e_j \rangle. \quad (2.2.8)$$

Equipping  $\text{End}(\mathcal{H})$  with the Hilbert-Schmidt inner product makes it into a complex  $d^2$ -dimensional Hilbert space.

**Definition 2.2.6.** Let  $d \in \mathbb{N}$  be a positive integer and let  $\mathcal{H}$  be a complex  $d$ -dimensional Hilbert space with norm  $\|\cdot\|$ . The **unit sphere** for  $\mathcal{H}$ , denoted  $S^{2d-1}$ , is the  $(2d - 1)$  dimensional manifold of vectors  $v \in \mathcal{H}$  with unit norm:

$$S^{2d-1} = \{v \in \mathcal{H} \mid \|v\| = 1\}. \quad (2.2.9)$$

**Definition 2.2.7.** Let  $\mathcal{H}$  be a complex finite-dimensional Hilbert space. A one-dimensional subspace of  $\mathcal{H}$  is called a *ray*. The *projective space* associated to  $\mathcal{H}$ , denoted by  $\mathbb{P}\mathcal{H}$ , is the set of rays in  $\mathcal{H}$ :

$$\mathbb{P}\mathcal{H} := \{\psi \subseteq \mathcal{H} \mid \dim(\psi) = 1\}. \quad (2.2.10)$$

Let the set of non-zero vectors in  $\mathcal{H}$  be denoted by

$$\mathcal{H}_\times = \mathcal{H} \setminus \{0\}. \quad (2.2.11)$$

Given a non-zero vector  $v \in \mathcal{H}_\times$ , the unique ray spanned by complex scalar multiples of  $v$  is written as

$$[v] := \mathbb{C}v = \{zv \in \mathcal{H} \mid z \in \mathbb{C}\} \in \mathbb{P}\mathcal{H}. \quad (2.2.12)$$

**Remark 2.2.8.** If the dimension of the complex Hilbert space  $\mathcal{H}$  is  $d = \dim(\mathcal{H}) \in \mathbb{N}$ , the projective space associated to  $\mathcal{H}$ ,  $\mathbb{P}\mathcal{H} \cong \mathbb{P}\mathbb{C}^d$ , is also sometimes denoted by  $\mathbb{C}P^{d-1}$ .

In this thesis, a *pure* quantum state will refer to a one-dimensional subspace  $\psi \subseteq \mathcal{H}$  of a complex finite-dimensional Hilbert space  $\mathcal{H}$ . In this manner, the set of all pure quantum states is identified with the projective space  $\mathbb{P}\mathcal{H}$ . Equivalently, one could consider pure quantum states to be equivalence classes of unit vectors  $v \in \mathcal{H}$  related to each other through multiplication by a complex phase  $e^{i\theta}$  for some  $\theta \in [0, 2\pi)$ . Another equivalent way to view pure quantum states is through the injective map,

$$P : \mathbb{P}\mathcal{H} \rightarrow \text{End}(\mathcal{H}) \quad (2.2.13)$$

from  $\mathbb{P}\mathcal{H}$  to  $\text{End}(\mathcal{H})$  sending each ray  $\psi \in \mathbb{P}\mathcal{H}$  to the unique operator  $P_\psi \in \text{End}(\mathcal{H})$  that is Hermitian ( $P_\psi^* = P_\psi$ ), is idempotent/projective ( $P_\psi^2 = P_\psi$ ), and has  $\psi$  as its eigenspace with eigenvalue one.

Alternatively, one can define  $P_\psi$  by choosing from the ray  $\psi$  any non-zero element  $v \in \psi_\times$  (recall that  $\psi_\times = \psi \setminus \{0\}$ ) as a fiducial or representative element of the ray. Using this representative element, together with an arbitrary operator  $X \in \text{End}(\mathcal{H})$ , we have

$$\text{Tr}(P_\psi X) = \frac{\langle v, Xv \rangle}{\langle v, v \rangle}. \quad (2.2.14)$$

Traditionally, one picks the fiducial element  $v \in \psi_\times$  to have unit norm, i.e.  $\langle v, v \rangle = 1$ , so that the above formula simplifies further to  $\text{Tr}(P_\psi X) = \langle v, Xv \rangle$ . The fundamental relationship between rays,  $\psi \in \mathbb{P}\mathcal{H}$ , and representation elements,  $v \in \psi_\times$ , captured by [Eq. \(2.2.14\)](#), will be used heavily in [Chapter 4](#).

Furthermore, if a linear operator  $X \in \text{End}(\mathcal{H})$  is *invertible*, then its kernel is empty and therefore for any ray  $\psi \in \mathbb{P}\mathcal{H}$ ,

$$\ker(X) \cap \psi \neq \emptyset. \quad (2.2.15)$$

In this way, every invertible linear operator,  $X \in \text{End}(\mathcal{H})$ , lifts to an action on the projective space,  $\mathbb{P}\mathcal{H}$ , such that  $X \cdot \psi \in \mathbb{P}\mathcal{H}$  is defined as the ray  $X \cdot \psi = [Xv]$  where  $v$  is any non-zero vector  $v \in \psi_\times$ .

A more general class of quantum states that go beyond pure quantum states are those which can be described by density operators.

**Definition 2.2.9.** Given a complex finite-dimensional Hilbert space, a *density operator* is an operator  $\rho \in \text{End}(\mathcal{H})$  such that  $\rho$  is

- *positive semidefinite*:  $\rho^* = \rho$  and  $\rho \geq 0$ , i.e.,

$$\forall v \in \mathcal{H} : \langle v, \rho v \rangle \geq 0, \quad (2.2.16)$$

- and *normalized*: meaning  $\text{Tr}(\rho) = 1$ .

The set of all density operators on  $\mathcal{H}$  will be expressed as  $\mathcal{S}$ ,  $\mathcal{S}(\mathcal{H})$  or  $\mathcal{S}_{\mathcal{H}}$ , depending on the context.

Evidently, a rank-one projector operator  $P_{\psi}$  onto a ray  $\psi \in \mathbb{P}\mathcal{H}$ , which models *pure* quantum states, is an example of a density operator. Density operators which are not projective correspond to *mixed* quantum states.

Occasionally, we will make use of a more abstract notion of a quantum state as a linear, positive and normalized map from a  $C^*$ -algebra to  $\mathbb{C}$ .

**Definition 2.2.10.** A *Banach algebra*  $\mathcal{A}$  is an associative algebra that is complete as a metric space induced by the norm  $\|\cdot\| : \mathcal{A} \rightarrow \mathbb{R}_{\geq 0}$  which is required to be submultiplicative:

$$\forall X, Y \in \mathcal{A} : \|XY\| \leq \|X\| \|Y\|. \quad (2.2.17)$$

A  *$C^*$  algebra* is a Banach algebra  $\mathcal{A}$  equipped with an map  $*$  :  $\mathcal{A} \rightarrow \mathcal{A}$  such that

- for all  $X \in \mathcal{A}$ ,  $(X^*)^* = X$ ,
- for all  $X, Y \in \mathcal{A}$ ,  $(X + Y)^* = X^* + Y^*$  and  $(XY)^* = Y^*X^*$ ,
- for all  $z \in \mathbb{C}$  and  $X \in \mathcal{A}$ ,  $(\lambda X)^* = \lambda^* X^*$ , and
- for all  $X \in \mathcal{A}$ ,  $\|X^*X\| = \|X\|^2$ .

**Remark 2.2.11.** A  $C^*$  algebra  $\mathcal{A}$  is said to be *unital* if the underlying algebra has a unit element  $1_{\mathcal{A}} \in \mathcal{A}$  such that

$$\forall X \in \mathcal{A} : 1_{\mathcal{A}}X = X1_{\mathcal{A}} = X. \quad (2.2.18)$$

Henceforth it will be implicitly assumed that all  $C^*$  algebras are unital.

**Example 2.2.12.** A particularly concrete example of a  $C^*$ -algebra is the algebra  $\text{Mat}_d(\mathbb{C})$  of matrices on a complex  $d$ -dimensional normed vector space  $\mathbb{C}^d$  where the involution  $X \mapsto X^*$  is the conjugate transpose of the matrix  $X \in \text{Mat}_d(\mathbb{C})$ . To make  $\text{Mat}_d(\mathbb{C})$  into a  $C^*$ -algebra, the norm  $\|\cdot\| : \text{Mat}_d(\mathbb{C})$  can be any submultiplicative norm. For instance, the *Frobenius norm* (also known as the *Hilbert Schmidt* norm),

$$\|X\|_{\text{F}} := \sqrt{\text{tr}(X^*X)}, \quad (2.2.19)$$

and the *operator norm*

$$\|X\|_{\text{op}} := \sup\{\|Xv\| \mid v \in V, \|v\| = 1\}, \quad (2.2.20)$$

are two examples of submultiplicative norms. These two norms are related by the inequality  $\|X\|_{\text{op}} \leq \|X\|_{\text{F}}$  which holds for any  $X \in \text{Mat}_d(\mathbb{C})$ . Furthermore, if  $\mathcal{H}$  is a  $d$ -dimensional complex Hilbert space, then after fixing an orthonormal basis we have an isomorphism  $\text{End}(\mathcal{H}) \simeq \text{Mat}_d(\mathbb{C})$  as  $C^*$ -algebras.

**Definition 2.2.13.** A *state* of a  $C^*$ -algebra  $\mathcal{A}$  is a map  $\varphi : \mathcal{A} \rightarrow \mathbb{C}$  which is

i) *linear*:  $\varphi(X + Y) = \varphi(X) + \varphi(Y)$ , for all  $X, Y \in \mathcal{A}$ ,

ii) *positive* meaning  $\varphi(X^*X) \geq 0$ , for all  $X \in \mathcal{A}$ , and

iii) *normalized* meaning  $\varphi(1_{\mathcal{A}}) = 1$  where  $1_{\mathcal{A}}$  is the unit element of  $\mathcal{A}$  (see [Remark 2.2.11](#)).

**Example 2.2.14.** The states of a  $C^*$ -algebra ([Definition 2.2.13](#)) generalize the aforementioned notions of pure (or mixed) states of a finite-dimensional Hilbert space. To see this explicitly, let  $\mathcal{A} = \text{End}(\mathcal{H})$  be the  $C^*$ -algebra of operators onto a complex finite-dimensional Hilbert space  $\mathcal{H}$  outlined in [Example 2.2.12](#). Associated to any ray  $\psi \in \mathbb{P}\mathcal{H}$  there exists a state  $\varphi_{\psi} : \text{End}(\mathcal{H}) \rightarrow \mathbb{C}$  defined for  $X \in \text{End}(\mathcal{H})$  by

$$\varphi_{\psi}(X) = \text{Tr}(P_{\psi}X) = \frac{\langle v, Xv \rangle}{\langle v, v \rangle}. \quad (2.2.21)$$

More generally, associated to any density operator  $\rho \in \text{End}(\mathcal{H})$  there exists a state  $\varphi_{\rho} : \text{End}(\mathcal{H}) \rightarrow \mathbb{C}$  defined for all  $X \in \text{End}(\mathcal{H})$  by

$$\varphi_{\rho}(X) = \text{Tr}(\rho X). \quad (2.2.22)$$

In either case, if  $\varphi : \text{End}(\mathcal{H}) \rightarrow \mathbb{C}$  is a state, the following inequalities hold for all  $X \in \text{End}(\mathcal{H})$ :

$$|\varphi(X)| \leq \|X\|_{\text{op}} \leq \|X\|_{\text{F}}. \quad (2.2.23)$$

**Lemma 2.2.15.** Let  $\varphi : \mathcal{A} \rightarrow \mathbb{C}$  be positive linear map on a  $C^*$ -algebra  $\mathcal{A}$ . Furthermore, let  $(X_1, \dots, X_k) \subset \mathcal{A}$  be a  $k$ -tuple of elements of  $\mathcal{A}$ . Define the *Gram matrix*  $G \in \text{Mat}_k(\mathbb{C})$  to be the  $k \times k$  complex matrix with entries  $G_{i,j} \in \mathbb{C}$  defined by

$$G_{i,j} := \varphi(X_i^* X_j). \quad (2.2.24)$$

Then  $G \geq 0$  is a positive semidefinite matrix, i.e., for all  $v = (v_1, \dots, v_k) \in \mathbb{C}^k$ ,

$$v^* G v = \sum_{i,j} v_i^* \varphi(X_i^* X_j) v_j = \varphi(Y^* Y) \geq 0. \quad (2.2.25)$$

where  $Y := \sum_{j=1}^k v_j X_j \in \mathcal{A}$ .

**Corollary 2.2.16.** Let  $\varphi : \mathcal{A} \rightarrow \mathbb{C}$  be a positive functional of a  $C^*$ -algebra  $\mathcal{A}$ . Then for all  $X, Y \in \mathcal{A}$ ,

$$|\varphi(X^* Y)|^2 \leq \varphi(X^* X) \varphi(Y^* Y). \quad (2.2.26)$$

*Proof.* Let  $Z = X + \lambda Y$  for some  $\lambda \in \mathbb{C}$  then

$$\varphi(Z^* Z) = \varphi((X + \lambda Y)^*(X + \lambda Y)) = \varphi(X^* X) + |\lambda|^2 \varphi(Y^* Y) + 2\text{Re}\{\lambda \varphi(X^* Y)\}. \quad (2.2.27)$$

Without loss of generality, for any real  $r \in \mathbb{R}$ , one can pick  $\lambda$  such that  $\lambda \varphi(X^* Y) = r |\varphi(X^* Y)| \in \mathbb{R}$ . Then

$$0 \leq \varphi(Z^* Z) = \varphi(X^* X) + r^2 \varphi(Y^* Y) + 2r |\varphi(X^* Y)|. \quad (2.2.28)$$

Since the right-hand side is a quadratic polynomial in  $r \in \mathbb{R}$ , the lower-bound above implies its discriminant must be non-positive

$$(2|\varphi(X^*Y)|)^2 - 4\varphi(Y^*Y)\varphi(X^*X) \leq 0, \quad (2.2.29)$$

which proves the lemma. Alternatively, the lemma follows by noticing that the difference  $\varphi(X^*X)\varphi(Y^*Y) - |\varphi(X^*Y)|^2$  is the determinant of the positive semidefinite matrix defined in [Lemma 2.2.15](#) for the case where  $k = 2$  and  $(X_1, X_2) = (X, Y)$ .  $\square$

## 2.2.2 Measurements

The most general form of quantum measurement that will be considered in this thesis is that of a positive-operator-valued measure, often abbreviated as POVM.

**Definition 2.2.17.** Let  $(X, \Sigma)$  be a measurable space, and let  $\text{End}(\mathcal{H})$  be the set of (necessarily bounded) operators acting on a finite-dimensional complex Hilbert space  $\mathcal{H}$ . Then a function  $E : \Sigma \rightarrow \mathcal{B}(\mathcal{H})$  is called a **positive-operator-valued measure** or **POVM** if it satisfies

- *positivity*: for all  $\Delta \in \Sigma$ , the operator  $E(\Delta)$  is positive semidefinite,

$$\forall \Delta \in \Sigma : E(\Delta) \geq 0, \quad (2.2.30)$$

- *nullity*:  $E(\emptyset) = 0_{\mathcal{H}}$  where  $0_{\mathcal{H}}$  is the zero operator on  $\mathcal{H}$ ,
- *countable additivity*: for all countable collections  $\{\Delta_j \in \Sigma\}_{j \in \mathbb{N}}$ , of pairwise disjoint sets,  $E$  satisfies

$$E\left(\bigcup_{j \in \mathbb{N}} \Delta_j\right) = \sum_{j \in \mathbb{N}} E(\Delta_j), \quad (2.2.31)$$

- *normalization*:  $E(X) = \mathbb{1}_{\mathcal{H}}$  where  $\mathbb{1}_{\mathcal{H}}$  is the identity operator on  $\mathcal{H}$ .

**Remark 2.2.18.** Note that a positive-operator-valued measure  $E : \Sigma(X) \rightarrow \mathcal{B}(\mathcal{H})$  can be viewed as a kind of state-dependent probability measure on  $X$ . More precisely, for each state  $\rho \in \mathcal{S}(\mathcal{H})$ , the map  $\xi_{\rho} : \Sigma(X) \rightarrow [0, 1]$  defined for  $\Delta \in \Sigma(X)$  by

$$\xi_{\rho}(\Delta) = \text{Tr}(\rho E(\Delta)) \quad (2.2.32)$$

is a probability measure on  $X$  in the sense of [Definition 2.1.3](#).

**Example 2.2.19.** Associated to every Hermitian operator  $A$  acting on a complex finite-dimensional complex Hilbert space  $\mathcal{H}$  is a canonical positive-operator-valued measure over the real numbers  $\mathbb{R}$  (equipped with the standard Borel  $\sigma$ -algebra on  $\mathbb{R}$ ), denoted by  $E_A : \Sigma(\mathbb{R}) \rightarrow \mathcal{B}(\mathcal{H})$ . Let  $\text{spec}(A) \subset \mathbb{R}$  be the spectrum of  $A$  (i.e., the set of its eigenvalues), and let

$$A = \sum_{a \in \text{spec}(A)} a P_a \quad (2.2.33)$$

be the spectral decomposition  $A$ , where  $P_a$  is the orthogonal projection operator onto eigenspace of  $A$  with eigenvalue  $a$ . Now for each interval  $\Delta \subset \mathbb{R}$  in  $\mathcal{B}(\mathcal{H})$ ,  $E_A(\Delta)$  is defined by

$$E_A(\Delta) = \sum_{a \in \Delta \cap \text{spec}(A)} P_a. \quad (2.2.34)$$

Indeed, for any quantum state  $\rho \in \mathcal{S}(\mathcal{H})$ , the function sending  $a \in \text{spec}(A)$  to  $p(a|\rho) \in [0, 1]$ , defined by

$$p(a|\rho) = \text{Tr}(\rho P_a), \quad (2.2.35)$$

is a discrete probability distribution over the spectrum of  $A$ .

**Example 2.2.20.** In most instances, a POVM  $E : \Sigma(X) \rightarrow \mathcal{B}(\mathcal{H})$  over a standard measurable space  $(X, \Sigma(X))$  can be constructed from a probability measure  $\nu$  on  $X$  together with a function  $g : X \rightarrow \mathcal{B}(\mathcal{H})$ , sometimes called the *positive-operator density*. Specifically, for  $\Delta \in \Sigma(X)$ , the value of  $E(\Delta)$  is expressed as

$$E(\Delta) := \int_{\Delta} g \, d\nu(x) = \int_{x \in \Delta} g(x) \, d\nu(x). \quad (2.2.36)$$

Of course, in order for this construction to be well-defined, the function  $g : X \rightarrow \mathcal{B}(\mathcal{H})$  needs to satisfy a number of conditions. First and foremost,  $g$  must be measurable (so that the integration above can be performed) and normalized such that  $E(X) = \mathbb{1}_{\mathcal{H}}$ . Moreover, to ensure the positivity condition holds, the function  $g : X \rightarrow \mathcal{B}(\mathcal{H})$  must, at the very least, be  $\nu$ -almost everywhere positive meaning there exists a sufficiently large (possibly-empty) subset  $N \in \Sigma(X)$  with zero measure,  $\nu(N) = 0$ , such that for all  $x \in X \setminus N$ ,  $g(x)$  is a positive-semidefinite operator.

## 2.3 Group theory

This section and the subsequent section (Section 2.4) reviews the essential ingredients of group theory and representation theory respectively. The conceptual advantages offered by acknowledging the import of representation theory into the development of quantum theory can be dated back to 1930 and Weyl's textbook on the subject [Wey50] (later translated). Any reader interested in approaching the subject of quantum theory from the perspective of group representation theory is recommended to read the accessible yet comprehensive introductory textbook by Woit, Woit, and Bartolini [WWB17].

### 2.3.1 Groups

**Definition 2.3.1.** A *group* is a non-empty set  $G$  equipped with a binary operation,  $\cdot : G \times G \rightarrow G$ , called the *group multiplication* which satisfies three conditions:

- i) *associativity*: for all  $a, b, c \in G$ ,  $(a \cdot b) \cdot c = a \cdot (b \cdot c)$ ,
- ii) *identity element*: there exists  $e \in G$  such that  $e \cdot g = g = g \cdot e$  for all  $g \in G$ , and
- iii) *invertibility*: for all  $g \in G$ , there exists a  $g^{-1} \in G$  such that  $g \cdot g^{-1} = g^{-1} \cdot g = e$ .



The cardinality of  $G$ , denoted  $|G|$ , is also called the *order* of the group.

**Definition 2.3.2.** A *group homomorphism* from  $G$  to  $G'$  is a function  $\alpha : G \rightarrow G'$  such that for all  $g_1, g_2 \in G$ ,

$$\alpha(g_1 \cdot g_2) = \alpha(g_1) \cdot \alpha(g_2). \quad (2.3.1)$$

If a group homomorphism,  $\alpha : G \rightarrow G'$ , is invertible and its inverse  $\alpha^{-1} : G' \rightarrow G$  is also a group homomorphism from  $G'$  to  $G$ , then  $\alpha$  is called a *group isomorphism* and  $G$  and  $G'$  are said to be *isomorphic*.

**Definition 2.3.3.** A group  $H$  is said to be a *subgroup* of a group  $G$ , and written as  $H \subseteq G$ , if there exists an injective group homomorphism  $\alpha : H \rightarrow G$  from  $H$  to  $G$ . Typically,  $H$  is identified as a subset in  $G$  through its image under  $\alpha$  (in which case  $\alpha(h) = h$ ). A *left coset* of a subgroup  $H$  in  $G$  is a set of the form

$$gH := \{g \cdot \alpha(h) \mid h \in H\}, \quad (2.3.2)$$

for some  $g \in G$ . Similarly, a *right coset* of a subgroup  $H$  in  $G$  is a set of the form

$$Hg := \{\alpha(h) \cdot g \mid h \in H\}, \quad (2.3.3)$$

for some  $g \in G$ .

**Definition 2.3.4.** Let  $X$  be a set. The *symmetric group on  $X$* , denoted by  $S_X$ , is the set of all bijective functions mapping  $X$  to itself with group multiplication given by function composition, i.e.  $f \cdot g := f \circ g$ .

**Remark 2.3.5.** Consider two sets,  $X$  and  $Y$ , and their respective symmetric groups,  $S_X$  and  $S_Y$ . Note that  $S_X$  and  $S_Y$  are isomorphic as groups if and only if  $X$  and  $Y$  are isomorphic as sets meaning they have the same cardinality,  $|X| = |Y|$ . The *symmetric group over  $n$  symbols*, denoted by  $S_n$ , can be understood (up to isomorphism) as the symmetric group over any set of cardinality  $n$ , e.g., the first  $n$  positive integers  $[n] = \{1, \dots, n\}$ .

**Definition 2.3.6.** Let  $G$  be a group,  $X$  be a set and  $S_X$  the symmetric group on  $X$ . A *group action* of  $G$  on  $X$  is a group homomorphism  $\alpha : G \rightarrow S_X$  from  $G$  to  $S_X$ . The group action of  $G$  of an element  $x \in X$  is typically abbreviated, for  $g \in G$ , by

$$g \cdot x := \alpha(g)(x). \quad (2.3.4)$$

The *orbit* of an element  $x \in X$  under the action of  $G$ , denoted by  $G \cdot x$  (or sometimes  $\alpha(G)(x)$ ), is the subset

$$G \cdot x := \{\alpha(g)(x) \in X \mid g \in G\}. \quad (2.3.5)$$

An element  $x \in X$  is said to be *invariant* if its orbit is a singleton, i.e.,  $G \cdot x = \{x\}$ .

**Example 2.3.7.** The set of all real numbers,  $\mathbb{R}$ , equipped with the binary operation of addition,  $+$  :  $\mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ , forms a group, called the *real additive group* where the identity element is zero and the inverse of  $r \in \mathbb{R}$  is the negation  $-r \in \mathbb{R}$ . The subset of all integers  $\mathbb{Z} = \{\dots, -2, -1, 0, +1, +2, \dots\} \subset \mathbb{R}$  is a subgroup of the real additive group  $\mathbb{R}$ . The set of non-zero real numbers, denoted by  $\mathbb{R}_\times = \mathbb{R} \setminus \{0\}$ , equipped with the binary operation of multiplication,  $\cdot$  :  $\mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ , forms a group, called the *real multiplicative group* where the identity element is one and the inverse of  $r \in \mathbb{R}_\times$  is  $\frac{1}{r} \in \mathbb{R}_\times$ . The finite subset  $\mathbb{Z}_2 = \{-1, +1\} \subset \mathbb{R}_\times$  is a subgroup of  $\mathbb{R}_\times$ . In an analogous fashion, the *complex additive group* is denoted  $\mathbb{C}$  and the *complex multiplicative group* is denoted by  $\mathbb{C}_\times$ .



## 2.3.2 Lie groups

There are many excellent textbook references for the subject of Lie group theory, e.g. [Wey46; Sim96; OV12; Cvi08; Pro07; Lee01]. Our primary reference for the theory of matrix Lie groups and their representations is the textbook by Hall [Hal15].

**Example 2.3.8.** Let  $\mathbb{F}$  be the field of complex, real or rational numbers, i.e.  $\mathbb{F} \in \{\mathbb{C}, \mathbb{R}, \mathbb{Q}\}$ . The set of  $d \times d$  matrices with entries in  $\mathbb{F}$  is denoted by  $\text{Mat}_d(\mathbb{F})$ . The *general linear group over  $\mathbb{F}$* , denoted by  $\text{GL}(d, \mathbb{F}) \subset \text{Mat}_d(\mathbb{F})$ , is the group of *invertible*  $d \times d$  matrices equipped with the binary operation of matrix multiplication, expressed using juxtaposition (i.e.  $(g_1, g_2) \mapsto g_1 g_2$ ), and identity element given by the identity matrix, denoted by either  $e$  or  $I$ . Throughout this thesis, our main examples will be the *real* general linear group,  $\text{GL}(d, \mathbb{R})$ , and the *complex* general linear group,  $\text{GL}(d, \mathbb{C})$ .

The set  $\text{Mat}_d(\mathbb{F})$  of  $d \times d$  matrices with entries in  $\mathbb{F}$  will always be considered as a topological space with respect to the standard topology on  $\text{Mat}_d(\mathbb{F})$  under the identification  $\text{Mat}_d(\mathbb{F}) \cong \mathbb{F}^{d \times d}$ . The subset of invertible matrices,  $\text{GL}(d, \mathbb{F})$ , and all of its subgroups will always inherit its topology from the standard topology on  $\text{Mat}_d(\mathbb{F})$ . The following definitions are concerned with topological properties of  $\text{GL}(d, \mathbb{C})$  and its subgroups.

**Definition 2.3.9.** A *linear group* is a subgroup  $G$  of  $\text{GL}(d, \mathbb{F})$  with binary operation given by matrix multiplication.

The following definition of a matrix Lie group is [Hal15, Defn. 1.4].

**Definition 2.3.10.** A *matrix Lie group*  $G$  is a linear group that is a subgroup of  $\text{GL}(d, \mathbb{C})$ , meaning every sequence  $\{g_k \in G \mid k \in \mathbb{N}\}$  of elements of  $G$  that converges to an element of  $\text{Mat}_d(\mathbb{C})$  either converges to an matrix in  $G$  or to a non-invertible matrix in  $\text{Mat}_d(\mathbb{C}) \setminus \text{GL}(d, \mathbb{C})$ .

**Example 2.3.11.** The real general linear group  $\text{GL}(d, \mathbb{R})$  and complex linear group  $\text{GL}(d, \mathbb{C})$  are both matrix Lie groups viewed as subgroups of  $\text{GL}(d, \mathbb{C})$ . Furthermore, the subgroups  $\text{SL}(d, \mathbb{R})$  and  $\text{SL}(d, \mathbb{C})$  of matrices with determinant one, known as the *special* linear groups, are matrix Lie groups.

**Example 2.3.12.** The *conjugate transpose* is an operation on complex matrices  $X \in \text{Mat}_d(\mathbb{C})$  sending the complex matrix  $X$  with entries  $X_{ij} \in \mathbb{C}$  to the complex matrix  $X^*$  with entries  $(X^*)_{ij} = X_{ji}^*$ . When restricted to the subgroup of real matrices  $\text{Mat}_d(\mathbb{R})$ , this operation is known as the *transpose* of a matrix and is denoted by  $X^T$  and has entries  $(X^T)_{ij} = X_{ji}$ . A complex matrix  $U \in \text{Mat}_d(\mathbb{C})$  is *unitary* if  $U^* = U^{-1}$ . Similarly, a real matrix  $O \in \text{Mat}_d(\mathbb{R})$  is *orthogonal* if  $O^* = O^{-1}$ . The subset of all unitary matrices in  $\text{Mat}_d(\mathbb{C})$  is a matrix Lie group called the *unitary group*  $U(d) \subset \text{GL}(d, \mathbb{C})$  and the subset of all orthogonal matrices in  $\text{Mat}_d(\mathbb{R})$  is a matrix Lie group called the *orthogonal group*  $O(d) \subset \text{GL}(d, \mathbb{R})$ . The *special unitary group* is  $SU(d) = U(d) \cap \text{SL}(d, \mathbb{C})$  and the *special orthogonal group* is  $SO(d) = O(d) \cap \text{SL}(d, \mathbb{R})$ .

The following definition combines [Hal15, Defn. 1.8] and the first part of [Hal15, Defn. 1.9].

**Definition 2.3.13.** A matrix Lie group  $G \subseteq \text{GL}(d, \mathbb{C})$  is said to be **compact** if it is compact as a subset of  $\text{Mat}_d(\mathbb{C}) \cong \mathbb{C}^{d \times d}$ . A matrix Lie group  $G \subseteq \text{GL}(d, \mathbb{C})$  is said to be **connected** if for any pair  $(g_1, g_2) \in G$ , there exists a *path* from  $g_1$  to  $g_2$ , which is a continuous function mapping  $t \in [0, 1]$  to  $g(t) \in G$  such that  $g(0) = g_1$  and  $g(1) = g_2$ .

By the Heine-Borel theorem, a matrix Lie group is compact if and only if it is i) closed as a subset of  $\text{Mat}_d(\mathbb{C})$ , meaning convergent sequences  $\{g_k \in G \mid k \in \mathbb{N}\}$  of matrices converge to an element of  $G$ , and ii) bounded, meaning there exists a finite  $C \in \mathbb{R}_{\geq 0}$  such that  $|g_{ij}| < C$  holds for all  $i, j \in [d]$  and  $g \in G$  [Hal15]. Also note that for general topological spaces, there is a distinction between the notions of connected and path-connected, but these notions are ultimately equivalent for matrix Lie groups [Hal15, Pg. 17].

**Example 2.3.14.** The groups  $\text{SU}(d), \text{U}(d), \text{SO}(d)$ , and  $\text{O}(d)$  (for all  $d \in \mathbb{N}$ ) are compact while the groups  $\text{GL}(d, \mathbb{C})$  and  $\text{SL}(d, \mathbb{C})$  (for  $d > 1$ ) are non-compact. The groups  $\text{SU}(d), \text{U}(d), \text{SO}(d), \text{GL}(d, \mathbb{C})$  and  $\text{SL}(d, \mathbb{C})$  (for all  $d \in \mathbb{N}$ ) are connected while  $\text{O}(d)$  is disconnected.

The advantage of restrictive our focus to matrix Lie groups, instead of the more general notion of non-matrix Lie groups, is simply that the exponential map  $\exp : \text{Mat}_d(\mathbb{C}) \rightarrow \text{Mat}_d(\mathbb{C})$  acting on complex matrices can be defined in a straight-forward manner and thus the Lie algebra of a matrix Lie group is also straight-forward to define.

**Definition 2.3.15.** The **exponential** of a matrix  $X \in \text{Mat}_d(\mathbb{C})$ , denoted by  $e^X$  or  $\exp(X)$  is the matrix defined by

$$\exp(X) := \sum_{n=0}^{\infty} \frac{X^n}{n!}. \quad (2.3.6)$$

**Remark 2.3.16.** A **one-parameter subgroup** of  $\text{GL}(d, \mathbb{C})$  ([Hal15, Defn. 2.13]) is a continuous group homomorphism  $A : \mathbb{R} \rightarrow \text{GL}(d, \mathbb{C})$ , meaning  $A(t)$  is a continuous function of  $t$ ,  $A(0) = I$  is the identity matrix and  $A(t_1 + t_2) = A(t_1)A(t_2)$  for all  $t_1, t_2 \in \mathbb{R}$ . For any complex matrix  $X \in \text{Mat}_d(\mathbb{C})$ , the function  $A(t) = \exp(tX)$  defines a one-parameter subgroup  $\text{GL}(d, \mathbb{C})$ . Moreover, by [Hal15, Thm. 2.14] every one-parameter subgroup of  $\text{GL}(d, \mathbb{C})$  is of this form for a *unique* complex matrix  $X \in \text{Mat}_d(\mathbb{C})$ .

### 2.3.3 Lie algebras

**Definition 2.3.17.** A **Lie algebra** is a vector space  $\mathfrak{g}$  over some field  $\mathbb{F}$  (e.g.,  $\mathbb{C}$  or  $\mathbb{R}$ ) equipped with a binary operation,  $[\cdot, \cdot] : \mathfrak{g} \times \mathfrak{g} \rightarrow \mathfrak{g}$ , called the **Lie bracket** which satisfies:

- *bilinearity*: for all  $a, b \in \mathbb{F}$  and  $X, Y, Z \in \mathfrak{g}$ ,

$$\begin{aligned} [aX + bY, Z] &= a[X, Z] + b[Y, Z], \\ [Z, aX + bY] &= a[Z, X] + b[Z, Y], \end{aligned} \quad (2.3.7)$$

- *alternativity*: for all  $X \in \mathfrak{g}$ ,

$$[X, X] = 0, \quad (2.3.8)$$

- and the *Jacobi identity*: for all  $X, Y, Z \in \mathfrak{g}$ ,

$$[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0. \quad (2.3.9)$$

All of the Lie algebras in this thesis will be Lie algebras associated to a matrix Lie group.

**Example 2.3.18.** Let  $G \subseteq \mathrm{GL}(d, \mathbb{C})$  be a matrix Lie group. The Lie algebra associated to  $G$ , denoted by  $\mathfrak{g}$ , is the subset of all matrices  $X \in \mathrm{Mat}_d(\mathbb{C})$  such that the image of the one-parameter subgroup  $A(t) = \exp(tX)$  is contained within  $G$ . In this case, the Lie bracket between the matrices  $X, Y \in \mathfrak{g}$  is given by the commutator of matrices,

$$[X, Y] = XY - YX. \quad (2.3.10)$$

**Example 2.3.19.** Perhaps the simplest non-trivial example of a Lie algebra is the Lie algebra  $\mathfrak{sl}(2, \mathbb{C})$  of  $2 \times 2$  complex matrices with trace zero. A standard basis for  $\mathfrak{sl}(2, \mathbb{C})$  are the matrices

$$H = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad X = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad Y = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad (2.3.11)$$

with commutation relations

$$[H, X] = 2X, \quad [H, Y] = 2Y, \quad [X, Y] = H. \quad (2.3.12)$$

The Lie algebra  $\mathfrak{sl}(2, \mathbb{C})$  is the Lie algebra associated to the Lie group  $\mathrm{SL}(2, \mathbb{C})$  of invertible  $2 \times 2$  complex matrices with unit determinant.

**Definition 2.3.20.** Let  $\mathfrak{h}$  and  $\mathfrak{g}$  be Lie algebras over the same field  $\mathbb{F}$ . A **Lie algebra homomorphism** from  $\mathfrak{h}$  to  $\mathfrak{g}$  is a linear map  $L : \mathfrak{h} \rightarrow \mathfrak{g}$  such that for all  $X, Y \in \mathfrak{h}$ ,

$$L([X, Y]) = [L(X), L(Y)]. \quad (2.3.13)$$

If  $L$  is additionally injective, meaning  $\ker(L) = 0$ , then  $\mathfrak{h}$  is a **Lie subalgebra** of  $\mathfrak{g}$ . Often Lie subalgebras are vector subspaces, i.e.,  $\mathfrak{h} \subseteq \mathfrak{g}$ .

**Example 2.3.21.** Let  $\mathfrak{g}$  be a Lie algebra with Lie bracket  $[\cdot, \cdot]$ . The **center** of  $\mathfrak{g}$ , denoted by  $\mathfrak{z}(\mathfrak{g})$ , is the subalgebra of elements  $X \in \mathfrak{g}$  which commute with all of  $\mathfrak{g}$ , i.e.,

$$\mathfrak{z}(\mathfrak{g}) := \{X \in \mathfrak{g} \mid \forall Y \in \mathfrak{g}, [X, Y] = 0\}. \quad (2.3.14)$$

If the Lie bracket vanishes everywhere, then  $\mathfrak{g} = \mathfrak{z}(\mathfrak{g})$  and  $\mathfrak{g}$  is said to be an **abelian** Lie algebra. The maximal abelian Lie subalgebra  $\mathfrak{h}$  of  $\mathfrak{g}$  is called the **Cartan subalgebra** of  $\mathfrak{g}$ .

**Remark 2.3.22.** If  $\mathcal{V}$  is a  $d$ -dimensional inner product space over the field  $\mathbb{F}$  with inner product  $\langle \cdot, \cdot \rangle$ , let  $\mathrm{GL}(\mathcal{V}) \subset \mathrm{End}(\mathcal{V})$  denote the group of invertible linear maps acting on  $\mathrm{GL}(\mathcal{V})$ . Given an orthonormal basis  $\{e_1, \dots, e_d\}$  on  $\mathcal{V}$ , we freely identify the group of invertible linear operators on  $\mathcal{V}$ ,  $\mathrm{GL}(\mathcal{V})$ , with the corresponding group of invertible matrices,  $\mathrm{GL}(d, \mathbb{F})$ , by defining the matrix  $M_g \in \mathrm{GL}(d, \mathbb{F})$  for  $g \in \mathrm{GL}(\mathcal{V})$  via

$$(M_g)_{i,j} = \langle e_i, g e_j \rangle. \quad (2.3.15)$$

In turn, the Lie algebra of  $\mathrm{GL}(\mathcal{V})$ , denoted by  $\mathfrak{gl}(\mathcal{V})$ , becomes identified with the set of all matrices  $\mathfrak{gl}(d, \mathbb{F}) = \mathrm{Mat}_d(\mathbb{F})$  with Lie bracket given by the matrix commutator. In a similar fashion, if the underlying field is the field of complex numbers,  $\mathbb{F} = \mathbb{C}$ , let  $\mathrm{U}(\mathcal{V}) \simeq \mathrm{U}(d)$  be the group of unitary maps acting on  $\mathcal{V}$  and if working over the real numbers,  $\mathbb{F} = \mathbb{R}$ , let  $\mathrm{O}(\mathcal{V}) \simeq \mathrm{O}(d)$  be the group of orthogonal maps acting on  $\mathcal{V}$ .

## 2.4 Representation theory

### 2.4.1 Group representations

The subject of representation theory, as a whole, is far too vast to provide a comprehensive introduction here. Instead, the primary focus of this section will be to introduce the reader to the representation theory of a rather special class of groups called reductive groups. The landmark result covered here will be the theorem of highest weights which provides helps to both classify and construct the representations of connected compact Lie groups. To begin we consider the definition of a representation and the notion of reducibility.

**Definition 2.4.1.** A **representation**,  $\Phi$ , of a group  $G$  on a vector space  $\mathcal{V}$  over the field  $\mathbb{F}$  is a group homomorphism,  $\Phi : G \rightarrow \text{GL}(\mathcal{V})$ , from the group  $G$  to the group of invertible linear operators on  $\mathcal{V}$ ,  $\text{GL}(\mathcal{V})$ , meaning

$$\forall g_1, g_2 \in G : \Phi(g_1 \cdot g_2) = \Phi(g_1)\Phi(g_2). \quad (2.4.1)$$

**Example 2.4.2.** Let  $n \in \mathbb{N}$  be a positive integer and let  $S_n$  be the symmetric group of permutations of order  $|S_n| = n!$ . For each vector space,  $\mathcal{V}$ , the **tensor-permutation representation** of  $S_n$  on  $\mathcal{V}^{\otimes n}$ , denoted by

$$T_n : S_n \rightarrow \text{GL}(\mathcal{V}^{\otimes n}), \quad (2.4.2)$$

is defined for each  $\sigma \in S_n$  and  $n$ -tuple of vectors  $(v_1, \dots, v_n) \in \mathcal{V}^{\times n}$ , by

$$T_n(\sigma)(v_1 \otimes \dots \otimes v_n) = v_{\sigma(1)} \otimes \dots \otimes v_{\sigma(n)}, \quad (2.4.3)$$

and extended linearly for all vectors in  $\mathcal{V}^{\otimes n}$ .

**Example 2.4.3.** Let  $\Phi : G \rightarrow \text{GL}(\mathcal{V})$  be a representation of a group  $G$  on a vector space  $\mathcal{V}$  over the field  $\mathbb{F}$  and let  $\mathcal{V}^*$  be the vector space dual to  $\mathcal{V}$  consisting of  $\mathbb{F}$ -valued linear functions on  $\mathcal{V}$ . The **dual representation** of  $G$  on  $\mathcal{V}^*$ , denoted by  $\Phi^* : G \rightarrow \text{GL}(\mathcal{V}^*)$ , is defined for all  $v \in \mathcal{V}$ ,  $g \in G$ , and  $f \in \mathcal{V}^*$  by

$$[\Phi^*(g)(f)](v) := f(\Phi(g^{-1})v). \quad (2.4.4)$$

**Example 2.4.4.** Let  $G \subseteq \text{GL}(n, \mathbb{C})$  be a matrix Lie group with Lie algebra  $\mathfrak{g}$ . The **adjoint representation** of  $G$  is the representation  $\text{Ad} : G \rightarrow \text{GL}(\mathfrak{g})$  defined for all  $X \in \mathfrak{g}$  by

$$\text{Ad}(g)(X) = \partial_{t=0}(ge^{tX}g^{-1}) = gXg^{-1}. \quad (2.4.5)$$

**Definition 2.4.5.** Let  $\Phi : G \rightarrow \text{GL}(\mathcal{V})$  be a representation of  $G$ . A subspace  $\mathcal{W} \subseteq \mathcal{V}$  is **invariant** if

$$\forall w \in \mathcal{W}, \forall g \in G : \Phi(g)w \in \mathcal{W}. \quad (2.4.6)$$

The **trivial** invariant subspaces of  $\mathcal{V}$  are the zero subspace  $0 \subseteq \mathcal{V}$  and the vector space  $\mathcal{V}$  itself. A representation is said to be **irreducible** if the only invariant subspaces of  $\mathcal{V}$  are the trivial ones.

**Definition 2.4.6.** Let  $\Phi : G \rightarrow \text{GL}(\mathcal{V})$  and  $\Phi' : G \rightarrow \text{GL}(\mathcal{V}')$  be representations of the group  $G$ . A linear operator  $L : \mathcal{V} \rightarrow \mathcal{V}'$  is said to be a **homomorphism** from  $\Phi$  to  $\Phi'$  (also called an *intertwining* or *equivariant* map) if

$$\forall g \in G : L\Phi(g) = \Phi'(g)L. \quad (2.4.7)$$

If  $L$  is additionally invertible, then it is called an **isomorphism** as its inverse,  $L^{-1} : \mathcal{V}' \rightarrow \mathcal{V}$ , is a homomorphism from  $\Phi'$  to  $\Phi$ .

Perhaps the most significant and widely applicable results in representation theory is Schur's lemma [Hal15; FH13].

**Lemma 2.4.7** (Schur's lemma). *Let  $\Phi : G \rightarrow \text{GL}(\mathcal{V})$  and  $\Phi' : G \rightarrow \text{GL}(\mathcal{V}')$  be irreducible representations of the group  $G$  over complex vector spaces and let  $L : \mathcal{V} \rightarrow \mathcal{V}'$  be an homomorphism from  $\Phi$  to  $\Phi'$ . Then either  $L$  is an isomorphism, or  $L$  is zero. Furthermore, if  $\mathcal{V} = \mathcal{V}'$ , then  $L$  is proportional to the identity,*

$$L = \alpha \mathbb{1}_{\mathcal{V}}. \quad (2.4.8)$$

for some  $\alpha \in \mathbb{C}$ .

*Proof.* As  $L$  is a homomorphism from  $\Phi$  to  $\Phi'$ , its kernel,

$$\ker(L) = \{v \in \mathcal{V} \mid Lv = 0 \in \mathcal{V}'\} \subseteq \mathcal{V}, \quad (2.4.9)$$

by Definition 2.4.6, is an invariant subspace of  $\mathcal{V}$  (because  $v \in \ker(L)$  implies  $L\Phi(g)v = \Phi'(g)Lv = 0$  and thus  $\Phi(g)v \in \ker(L)$ ). Since  $\Phi$  is irreducible,  $\ker(L)$  is either zero, in which case  $L$  is invertible and thus an isomorphism, or all of  $\mathcal{V}$ , in which case  $L$  is zero.

Finally, if  $\mathcal{V} = \mathcal{V}'$ , then for any eigenvalue  $\alpha \in \mathbb{C}$  of  $L$ , the homomorphism  $L - \alpha \mathbb{1}$  has non-empty kernel which means  $L - \alpha \mathbb{1}$  must be zero and thus  $L = \alpha \mathbb{1}$  as claimed.  $\square$

## 2.4.2 Lie algebra representations

It will also be useful to consider the representation theory of Lie algebras, which parallels Definition 2.4.1.

**Definition 2.4.8.** Let  $\mathfrak{g}$  be a Lie algebra with Lie bracket  $[\cdot, \cdot] : \mathfrak{g} \times \mathfrak{g} \rightarrow \mathfrak{g}$ , and let  $\mathfrak{gl}(\mathcal{V})$  be the Lie algebra of linear maps on a vector space  $\mathcal{V}$ . A **representation of a Lie algebra**  $\mathfrak{g}$  on  $\mathcal{V}$  is a Lie algebra homomorphism,  $L : \mathfrak{g} \rightarrow \mathfrak{gl}(\mathcal{V})$  (Definition 2.3.20), meaning

$$\forall X, Y \in \mathfrak{g} : L([X, Y]) = [L(X), L(Y)]. \quad (2.4.10)$$

**Lemma 2.4.9.** *Let  $G$  be a Lie group with Lie algebra  $\mathfrak{g}$  and let  $\Phi : G \rightarrow \text{GL}(\mathcal{V})$  be a representation of  $G$ . Define  $\phi : \mathfrak{g} \rightarrow \mathfrak{gl}(\mathcal{V})$  for all  $X \in \mathfrak{g}$  and  $v \in \mathcal{V}$  by*

$$\phi(X)v = \partial_{t=0}\Phi(\exp(tX))v. \quad (2.4.11)$$

Then  $\phi : \mathfrak{g} \rightarrow \mathfrak{gl}(\mathcal{V})$  is a representation of the Lie algebra  $\mathfrak{g}$  called the **induced representation**.

**Definition 2.4.10.** The *adjoint representation* of a Lie algebra  $\mathfrak{g}$  on itself,

$$\text{ad} : \mathfrak{g} \rightarrow \mathfrak{gl}(\mathfrak{g}), \quad (2.4.12)$$

is the representation of  $\mathfrak{g}$  defined, for  $X, Y \in \mathfrak{g}$  by

$$\text{ad}(X)Y = [X, Y]. \quad (2.4.13)$$

where  $[\cdot, \cdot]$  is the Lie bracket for  $\mathfrak{g}$ .

That the adjoint representation of  $\mathfrak{g}$  satisfies the definition of a Lie algebra representation, [Definition 2.4.8](#), follows directly from the Jacobi identity. Furthermore, if  $\mathfrak{g}$  is the Lie algebra of some Lie group  $G$ , then the Lie algebra representation of  $\mathfrak{g}$  induced by the adjoint representation of the Lie group  $G$  on  $\mathfrak{g}$  coincides with the above definition. For matrix Lie groups,  $G$ , this correspondence is especially easy to verify.

**Example 2.4.11.** Let  $\text{Ad} : G \rightarrow \text{GL}(\mathfrak{g})$  be the adjoint representation of the matrix Lie group  $G$  with Lie algebra  $\mathfrak{g}$ . The induced representation,  $\text{ad} : \mathfrak{g} \rightarrow \mathfrak{gl}(\mathfrak{g})$ , of  $\mathfrak{g}$  on itself is therefore

$$\text{ad}(X)Y = \partial_{t=0} \text{Ad}(\exp(tX))Y = \partial_{t=0}(e^{tX}Y e^{-tX}) = XY - YX = [X, Y]. \quad (2.4.14)$$

An important property of Lie groups  $G$  that are connected as topological groups is that the group  $G$  and its Lie algebra  $\mathfrak{g}$  share invariant subspaces.

**Lemma 2.4.12.** *Let  $\Phi : G \rightarrow \text{GL}(\mathcal{V})$  be a finite dimensional representation of a connected Lie group  $G$  and let  $\phi : \mathfrak{g} \rightarrow \mathfrak{gl}(\mathcal{V})$  be the induced representation of its Lie algebra. A subspace  $\mathcal{U} \subseteq \mathcal{V}$  is a  $\Phi$ -invariant if and only if it is  $\phi$ -invariant.*

*Proof.* If  $\mathcal{U}$  is  $\Phi$ -invariant, then for all  $t \in \mathbb{R}$  and  $X \in \mathfrak{g}$ ,  $\Phi(\exp(tX))v \in \mathcal{U}$  and thus  $\phi(X)v = \partial_{t=0} \Phi(\exp(tX))v \in \mathcal{U}$ . On the other hand, if  $\mathcal{U}$  is  $\phi$ -invariant, then for any  $X \in \mathfrak{g}$ ,

$$\Phi(\exp(X))v = \exp(\phi(X))v = \sum_{k=0}^{\infty} \frac{1}{k!} (\phi(X))^k v \in \mathcal{U}, \quad (2.4.15)$$

Since  $X \in \mathfrak{g}$  was arbitrary and a connected Lie group  $G$  is generated by elements of the form  $\exp(X)$  ([\[Hal15, Cor. 37\]](#)), the above equation implies that  $\mathcal{U}$  is also  $\Phi$  invariant.  $\square$

Our next example is concerned with deconstructing the irreducible representations of  $\mathfrak{sl}(2, \mathbb{C})$  and is a well-known construction in the representation theory of Lie algebras. Our presentation closely mirrors [\[Hal15, Sec. 4.2\]](#).

**Example 2.4.13.** Recall the Lie algebra,  $\mathfrak{sl}(2, \mathbb{C})$ , of  $2 \times 2$  complex, traceless matrices. Now suppose  $\phi : \mathfrak{sl}(2, \mathbb{C}) \rightarrow \mathfrak{gl}(d, \mathbb{C})$  is an irreducible  $d$ -dimensional representation of  $\mathfrak{sl}(2, \mathbb{C})$  and let  $H, X, Y \in \mathfrak{sl}(2, \mathbb{C})$  be the basis for  $\mathfrak{sl}(2, \mathbb{C})$  defined in [Example 2.3.19](#). Now let  $\mu \in \mathbb{C}$  be an eigenvector of  $\phi(H)$  with eigenvector  $v$ ,

$$\phi(H)v = \mu v. \quad (2.4.16)$$

Using the commutation relation  $[H, X] = 2X$  and the assumption that  $\phi$  preserves the Lie bracket, we conclude that  $\phi(X)v$  must either be the zero vector or an eigenvector of  $\phi(H)$  with eigenvalue  $\mu + 2$  because

$$\phi(H)\phi(X)v = \phi(X)\phi(H)v + \phi([H, X])v = (\mu + 2)\phi(X)v. \quad (2.4.17)$$

Since the representation is assumed finite,  $\phi(H)$  can only have finitely many eigenvalues and thus repeated application of  $\phi(X)$  to  $v$  must eventually produce a zero vector, i.e., there exists a positive integer  $r$  such that

$$u := \phi(X)^r v \neq 0, \quad \phi(X)^{r+1}v = 0. \quad (2.4.18)$$

Therefore,  $u$  is a non-zero eigenvector of  $\phi(H)$  with eigenvalue  $\lambda := \mu + 2r$ .

In a similar fashion to Eq. (2.4.17), as  $[H, Y] = -2Y$ , we conclude that  $\phi(Y)u_0$  is either zero or an eigenvector of  $\phi(H)$  with eigenvalue  $\lambda - 2$ . Again, by the finite-dimensionality of the representation, we conclude there exists another positive integer  $\ell$  such that

$$\phi(Y)^\ell u \neq 0, \quad \phi(Y)^{\ell+1}u = 0. \quad (2.4.19)$$

For each  $j \in \{0, \dots, \ell\}$ , let  $w_j$  be the non-zero eigenvector of  $\phi(H)$ ,

$$w_j := \phi(Y)^j u, \quad (2.4.20)$$

with eigenvalue  $\lambda - 2j$ , i.e.,

$$\phi(H)w_j = (\lambda - 2j)w_j. \quad (2.4.21)$$

Evidently, these eigenvectors span an  $(\ell + 1)$ -dimensional subspace that is invariant under the action of  $\mathfrak{sl}(2, \mathbb{C})$  via the representation  $\phi : \mathfrak{sl}(2, \mathbb{C}) \rightarrow \mathfrak{gl}(d, \mathbb{C})$ . Under the assumption that  $\phi$  is irreducible, we conclude  $d = \ell + 1$ .

Finally, we can characterize the action of  $\phi(X)$  on  $w_j$  for  $j > 0$  using induction and the commutation relations  $[X, Y] = H$  to obtain

$$\phi(X)w_j = j(\lambda - (j - 1))w_{j-1}. \quad (2.4.22)$$

By applying  $\phi(Y)$  to both sides of this expression, and using the fact that  $\phi(Y)w_\ell = 0$ , one obtains

$$0 = (\ell + 1)(\lambda - \ell)w_\ell, \quad (2.4.23)$$

and thus concludes that the eigenvalue  $\lambda$  must equal the integer  $\ell$ .

A key tool for understanding the anatomy of a Lie algebra representation,  $\phi : \mathfrak{g} \rightarrow \mathfrak{gl}(\mathcal{V})$  is to examine its restriction onto a subalgebra which is easier to study. Specifically, given any abelian subalgebra  $\mathfrak{h} \subseteq \mathfrak{g}$ , the fact that  $[X, Y] = 0$  for all  $X, Y \in \mathfrak{h}$  implies that the operators  $\phi(X)$  and  $\phi(Y)$  acting on  $\mathcal{V}$  must commute.

**Definition 2.4.14.** Let  $\phi : \mathfrak{g} \rightarrow \mathfrak{gl}(\mathcal{V})$  be a Lie algebra representation and fix  $\mathfrak{h} \subseteq \mathfrak{g}$  a maximal abelian subalgebra of  $\mathfrak{g}$ . A *weight* of the representation is a linear function,  $\mu \in \mathfrak{h}^*$ , such that there exists a non-zero vector,  $v \in \mathcal{V}$ , called a *weight vector*, satisfying

$$\forall H \in \mathfrak{h} : \phi(H)v = \mu(H)v. \quad (2.4.24)$$

The subspace of all weight vectors for a given weight  $\mu \in \mathfrak{h}^*$ , denoted by  $\mathcal{V}_\mu \subseteq \mathcal{V}$ , is called the *weight space* and its dimension is the *multiplicity* of  $\mu$ .



**Example 2.4.15.** Perhaps the most important representation of a Lie algebra  $\mathfrak{g}$  is the adjoint representation,  $\text{ad} : \mathfrak{g} \rightarrow \mathfrak{gl}(\mathfrak{g})$ , of  $\mathfrak{g}$  on itself. Once a maximal abelian subalgebra  $\mathfrak{h} \subseteq \mathfrak{g}$  is fixed, the weights, weight vectors, and weight spaces associated to the adjoint representation are called roots, root vectors, and root spaces, respectively. A **root** is a linear function  $\alpha \in \mathfrak{h}^*$  such that there exists a non-zero **root vector**  $X \in \mathfrak{g}$  satisfying

$$\forall H \in \mathfrak{h} : \text{ad}(H)X = [H, X] = \alpha(H)X. \quad (2.4.25)$$

The **root space** of all root vectors for the root  $\alpha$  is denoted by  $\mathfrak{g}_\alpha \subseteq \mathfrak{g}$ .

### 2.4.3 Composition

Before discussing how to construct a representation from scratch, it will be helpful to understand a few of the operations that can be used to compose representations to produce new ones. Note that vector spaces in this section are considered to be over the same field, e.g., the complex numbers  $\mathbb{C}$ .

The first two operations combine two representations of the *same* group, albeit on possibly different vector spaces.

**Definition 2.4.16.** Let  $\Phi_1 : G \rightarrow \text{GL}(\mathcal{V}_1)$  and  $\Phi_2 : G \rightarrow \text{GL}(\mathcal{V}_2)$  be representations of the group  $G$ . The **internal direct sum** of  $\Phi_1$  and  $\Phi_2$  is a representation of the form

$$\Phi_1 \oplus \Phi_2 : G \rightarrow \text{GL}(\mathcal{V}_1 \oplus \mathcal{V}_2) \quad (2.4.26)$$

and is defined, for all  $g \in G$ , by

$$(\Phi_1 \oplus \Phi_2)(g) = \Phi_1(g) \oplus \Phi_2(g). \quad (2.4.27)$$

**Definition 2.4.17.** Let  $\Phi_1 : G \rightarrow \text{GL}(\mathcal{V}_1)$  and  $\Phi_2 : G \rightarrow \text{GL}(\mathcal{V}_2)$  be representations of the group  $G$ . The **internal tensor product** of  $\Phi_1$  and  $\Phi_2$  is a representation of the form

$$\Phi_1 \otimes \Phi_2 : G \rightarrow \text{GL}(\mathcal{V}_1 \otimes \mathcal{V}_2) \quad (2.4.28)$$

and is defined, for all  $g \in G$ , by

$$(\Phi_1 \otimes \Phi_2)(g) = \Phi_1(g) \otimes \Phi_2(g). \quad (2.4.29)$$

The next two operations combine two representations of potentially *distinct* groups.

**Definition 2.4.18.** Let  $\Phi_1 : G_1 \rightarrow \text{GL}(\mathcal{V}_1)$  and  $\Phi_2 : G_2 \rightarrow \text{GL}(\mathcal{V}_2)$  be representations of the groups  $G_1$  and  $G_2$ . The **external direct sum** of  $\Phi_1$  and  $\Phi_2$  is a representation of the form

$$\Phi_1 \boxplus \Phi_2 : G_1 \times G_2 \rightarrow \text{GL}(\mathcal{V}_1 \oplus \mathcal{V}_2) \quad (2.4.30)$$

and is defined, for all  $g_1 \in G_1$  and  $g_2 \in G_2$ , by

$$(\Phi_1 \boxplus \Phi_2)(g_1, g_2) = \Phi_1(g_1) \oplus \Phi_2(g_2). \quad (2.4.31)$$



**Definition 2.4.19.** Let  $\Phi_1 : G_1 \rightarrow \text{GL}(\mathcal{V}_1)$  and  $\Phi_2 : G_2 \rightarrow \text{GL}(\mathcal{V}_2)$  be representations of the groups  $G_1$  and  $G_2$ . The **external tensor product** of  $\Phi_1$  and  $\Phi_2$  is a representation of the form

$$\Phi_1 \boxtimes \Phi_2 : G_1 \times G_2 \rightarrow \text{GL}(\mathcal{V}_1 \otimes \mathcal{V}_2) \quad (2.4.32)$$

and is defined, for all  $g_1 \in G_1$  and  $g_2 \in G_2$ , by

$$(\Phi_1 \boxtimes \Phi_2)(g_1, g_2) = \Phi_1(g_1) \otimes \Phi_2(g_2). \quad (2.4.33)$$

**Remark 2.4.20.** Note there is seldom an opportunity to confuse the internal and external constructions when  $G_1$  and  $G_2$  are distinct groups as only the external construction is defined in this case. If however, the two groups are identical, i.e.  $G_1 = G_2 = G$ , then the internal and external constructions can be related using the diagonal (or copy) embedding:

$$\Delta : G \rightarrow G \times G :: g \mapsto (g, g). \quad (2.4.34)$$

In such cases, the internal and external constructions of the representations  $\Phi_1 : G \rightarrow \text{GL}(\mathcal{V}_1)$  and  $\Phi_2 : G \rightarrow \text{GL}(\mathcal{V}_2)$  are related by the formulas

$$\Phi_1 \oplus \Phi_2 = (\Phi_1 \boxplus \Phi_2) \circ \Delta, \quad (2.4.35)$$

$$\Phi_1 \otimes \Phi_2 = (\Phi_1 \boxtimes \Phi_2) \circ \Delta. \quad (2.4.36)$$

For the sake of brevity, the *tensor product* of two representations will always refer to the *internal* tensor product. Similarly, the *direct sum* of two representations will always refer to the *internal* direct sum.

Under certain circumstances, there is an opportunity to combine representations acting on the *same* vector space.

**Definition 2.4.21.** Let  $\Phi_1 : G_1 \rightarrow \text{GL}(\mathcal{V})$  and  $\Phi_2 : G_2 \rightarrow \text{GL}(\mathcal{V})$  be representations of  $G_1$  and  $G_2$  on a common vector space  $\mathcal{V}$ . The representations  $\Phi_1$  and  $\Phi_2$  are said to be **mutually commuting** representations on  $\mathcal{V}$  if

$$\forall g_1 \in G_1, g_2 \in G_2 : [\Phi_1(g_1), \Phi_2(g_2)] = 0. \quad (2.4.37)$$

If  $\Phi_1$  and  $\Phi_2$  are mutually commuting, then one can construct the representation

$$\Phi_{1,2} : G_1 \times G_2 \rightarrow \text{GL}(\mathcal{V}) \quad (2.4.38)$$

defined, for all  $g_1 \in G_1$  and  $g_2 \in G_2$ , by

$$(\Phi_{1,2})(g_1, g_2) = \Phi_1(g_1)\Phi_2(g_2) = \Phi_1(g_1)\Phi_2(g_2). \quad (2.4.39)$$

**Example 2.4.22.** Given a representation  $\Phi : G \rightarrow \text{GL}(\mathcal{V})$  and a positive integer  $n \in \mathbb{N}$ , one can construct a representation of  $G$  on  $\mathcal{V}^{\otimes n}$  called the ***n*th tensor power representation** of  $\Phi$ , denoted  $\Phi^{\otimes n} : G \rightarrow \text{GL}(\mathcal{V}^{\otimes n})$ , defined as the internal tensor product of  $n$  copies of  $\Phi$ . A related representation is obtained by first noting that  $\Phi^{\otimes n}(g) = \Phi(g)^{\otimes n}$  commutes with the permutation of the  $n$  tensor factors in  $\mathcal{V}^{\otimes n}$ . Then let  $\iota_n : \text{Sym}^n(\mathcal{V}) \hookrightarrow \mathcal{V}^{\otimes n}$  be the  $\text{GL}(\mathcal{V})$ -equivariant isometry from the symmetric subspace of degree  $n$ ,  $\text{Sym}^n(\mathcal{V})$ , to  $\mathcal{V}^{\otimes n}$ . Finally one can define the ***n*th symmetric power representation**, denoted  $\Phi^{\vee n} : G \rightarrow \text{GL}(\text{Sym}^n(\mathcal{V}))$ , for all  $g \in G$  by

$$\Phi^{\vee n}(g) := \iota_n^*(\Phi(g))^{\otimes n} \iota_n. \quad (2.4.40)$$

Analogously, if  $v \in \mathcal{V}$  and  $n \in \mathbb{N}$ , define  $v^{\vee n} \in \text{Sym}^n \mathcal{V}$  by  $v^{\vee n} := \iota_n^* v^{\otimes n}$ .

## 2.4.4 Decomposition

**Definition 2.4.23.** A representation  $\Phi : G \rightarrow \text{GL}(\mathcal{V})$  is said to be *completely reducible* if it is isomorphic to a finite direct sum of irreducible representations, i.e., there exists a finite set of irreducible representations  $\Phi_i : G \rightarrow \text{GL}(\mathcal{V}_i)$  indexed by  $i \in [r]$  such that  $\mathcal{V}$  decomposes as a vector space into a finite direct sum,

$$\mathcal{V} = \mathcal{V}_1 \oplus \mathcal{V}_2 \cdots \oplus \mathcal{V}_r, \quad (2.4.41)$$

and furthermore  $\Phi(g)$  decomposes, for each  $g$ , as an operator

$$\Phi(g) = \Phi_1(g) \oplus \cdots \oplus \Phi_r(g). \quad (2.4.42)$$

The group  $G$  itself is said to be *completely reducible* if every finite-dimensional representation of  $G$  is completely reducible.

The following is [Hal15, Prop. 4.27].

**Lemma 2.4.24.** *Let  $\Phi : K \rightarrow \text{U}(\mathcal{V})$  be a finite-dimensional unitary representation of  $K$  on a Hilbert space  $\mathcal{V}$  with inner product  $\langle \cdot, \cdot \rangle$ . Then  $\Phi$  is completely reducible.*

*Proof.* First, we prove that any reducible, unitary representation  $\Phi : K \rightarrow \text{U}(\mathcal{V})$  of  $K$  on the Hilbert space  $\mathcal{V}$  splits into the orthogonal direct sum of unitary representations. Let  $\mathcal{W} \subset \mathcal{V}$  be a proper invariant subspace of  $\mathcal{V}$ . The orthogonal complement  $\mathcal{W}^\perp$  of  $\mathcal{W}$  (evaluated with respect to the inner product  $\langle \cdot, \cdot \rangle$  on  $\mathcal{V}$ ), leads to the decomposition  $\mathcal{V} = \mathcal{W} \oplus \mathcal{W}^\perp$ . That  $\mathcal{W}^\perp$  is also an invariant subspace follows from the unitarity of  $\Phi(k)$  for all  $k \in K$ . The invariance of  $\mathcal{W}$  means  $\Phi(k)w \in \mathcal{W}$  for all  $w \in \mathcal{W}$  and  $k \in K$  (including  $k^{-1}$ ). Therefore, for generic  $w' \in \mathcal{W}^\perp$  and  $w \in \mathcal{W}$ ,

$$0 = \langle w', \Phi(k^{-1})w \rangle = \langle w', \Phi(k)^*w \rangle = \langle \Phi(k)w', w \rangle, \quad (2.4.43)$$

which means  $\Phi(k)w' \in \mathcal{W}^\perp$  and thus  $\mathcal{W}^\perp$  is also invariant.

Second, if the representation on  $\mathcal{W}$  (or  $\mathcal{W}^\perp$ ) is reducible, then the above argument can be applied to recursively split the subrepresentation on  $\mathcal{W}$  (or  $\mathcal{W}^\perp$ ). Since  $\mathcal{V}$  is finite-dimensional and one-dimensional representations are irreducible by definition, this recursive process must terminate, producing an orthogonal direct-sum decomposition of  $\mathcal{V}$  into irreducible invariant subspaces.  $\square$

The following is [Hal15, Thm 4.28] as is sometimes known as *Weyl's unitary trick*.

**Theorem 2.4.25.** *If  $K$  is a compact matrix Lie group, then every finite-dimensional representation of  $K$  is completely reducible.*

*Proof.* Let  $\Phi : K \rightarrow \text{GL}(\mathcal{V})$  be a representation of  $K$  on a finite-dimensional Hilbert space  $\mathcal{V}$  with inner product  $\langle \cdot, \cdot \rangle$ . If  $\Phi$  is a unitary representation, i.e.  $\Phi(k)^* = \Phi(k^{-1})$  (with respect to  $\langle \cdot, \cdot \rangle$ ) for all  $k \in K$ , then Lemma 2.4.24 implies  $\Phi$  is completely reducible. Thus the only remaining case to consider is when  $\Phi$  is not a unitary representation. The trick here is to show that whenever  $\Phi(k)$  is not unitary with respect to the inner product  $\langle \cdot, \cdot \rangle$ , one can

define a new inner product,  $\langle \cdot, \cdot \rangle_K$  on  $\mathcal{V}$  such that  $\Phi(k)$  is unitary. This new inner product is obtained by *twirling* the old inner product, namely for  $v, w \in \mathcal{V}$ ,

$$\langle v, w \rangle_G := \int_G \langle \Phi(k)v, \Phi(k)w \rangle d\mu(k), \quad (2.4.44)$$

where  $d\mu$  is the unique up to scaling left and right  $K$ -invariant Haar measure on  $K$  (compactness of  $K$  ensures this existence and invariance of this measure, as well as the convergence of the integral). That  $\Phi(k)$  is a unitary operator with respect to  $\langle \cdot, \cdot \rangle_K$  follows from the invariance of the Haar measure  $d\mu$  which implies  $\langle \Phi(k)v, \Phi(k)w \rangle_K = \langle v, w \rangle_K$  for all  $v, w \in \mathcal{V}$  which means  $\Phi(k)^* \Phi(k) = \mathbb{1}_{\mathcal{V}}$  is the identity on  $\mathcal{V}$  and thus  $\Phi(k)^{-1} = \Phi(k^{-1}) = \Phi(k)^*$  is unitary for all  $g \in G$ . By [Lemma 2.4.24](#), again,  $\Phi$  is completely reducible.  $\square$

## 2.4.5 Complexification & semisimplicity

At this stage, we turn our attention to a particular class of Lie algebras called *semisimple Lie algebras* which are obtained from the Lie algebras of compact Lie groups by a process known as *complexification*. Our interest in semisimple Lie algebras stems from the fact that, much like [Theorem 2.4.25](#), their finite-dimensional representations are completely reducible [[Hal15](#), Thm. 10.9].

Semisimple Lie algebras include many of the familiar examples of including  $\mathfrak{sl}(d, \mathbb{C})$ ,  $\mathfrak{so}(d, \mathbb{C})$ , and  $\mathfrak{sp}(d, \mathbb{C})$ . Our main goal will be to understand the decomposition of reductive Lie algebras and establish some consistent notation along the way.

To begin, we consider the complexification of a real Lie algebras,  $\mathfrak{k}$ , which is formed by extending the field of real numbers,  $\mathbb{R}$ , to the field of complex numbers,  $\mathbb{C}$ , in the usual way.

**Definition 2.4.26.** Let  $\mathfrak{k}$  be a Lie algebra with Lie bracket  $[\cdot, \cdot]$ . The *complexification* of  $\mathfrak{k}$ , expressed as  $\mathfrak{k}_{\mathbb{C}}$  or  $\mathfrak{k} \oplus i\mathfrak{k}$ , is the Lie algebra of formal linear combinations,

$$X = X_1 + iX_2, \quad (2.4.45)$$

where  $X_1, X_2 \in \mathfrak{k}$  and where  $iX := -X_2 + iX_1$ . The Lie bracket is extended linearly for  $X, Y \in \mathfrak{k}_{\mathbb{C}}$  to

$$[X, Y] = ([X_1, Y_1] - [X_2, Y_2]) + i([X_1, Y_2] + [X_2, Y_1]). \quad (2.4.46)$$

**Example 2.4.27.** The complexification of the Lie algebra of real  $d \times d$  matrices is isomorphic to the Lie algebra of complex  $d \times d$  matrices,

$$\mathfrak{gl}(d, \mathbb{R})_{\mathbb{C}} \cong \mathfrak{gl}(d, \mathbb{C}). \quad (2.4.47)$$

Similarly, the complexification of the Lie algebra of skew-Hermitian  $d \times d$  matrices (which is a real Lie algebra) is also isomorphic to the Lie algebra of complex  $d \times d$  matrices,

$$\mathfrak{u}(d)_{\mathbb{C}} \cong \mathfrak{gl}(d, \mathbb{C}). \quad (2.4.48)$$

This second example follows because every  $d \times d$  complex matrix  $M \in \mathfrak{gl}(d, \mathbb{C})$  is expressible as

$$M = \left( \frac{M - M^*}{2} \right) + i \left( \frac{M + M^*}{2i} \right) \quad (2.4.49)$$

where  $M^*$  is the conjugate transpose of the matrix  $M$  and where *both* the first and second terms in parentheses are skew-Hermitian matrices, and thus elements of  $\mathfrak{u}(d)_{\mathbb{C}}$ . For more examples of complexifications, see [Hal15, Eq. 3.17].

**Remark 2.4.28.** When working with Lie algebras and their complexifications, it can quickly become quite challenging to keep track of all of the different relationships between the various subspaces. It is therefore essential to establish some consistent notational conventions. Our notational conventions are fairly standard and strongly mirrors those of Ref. [BCV21]. As a starting point we consider a real Lie algebra  $\mathfrak{k}$  and its complexification

$$\mathfrak{g} = \mathfrak{k}_{\mathbb{C}} = \mathfrak{k} \oplus i\mathfrak{k}. \quad (2.4.50)$$

Acting on the complexified Lie algebra  $\mathfrak{g} = \mathfrak{k}_{\mathbb{C}}$ , there exists a map, called the *local Cartan involution*,

$$\theta : \mathfrak{g} \rightarrow \mathfrak{g}, \quad (2.4.51)$$

which sends each  $X = X_1 + iX_2 \in \mathfrak{g}$  (where  $X_1, X_2 \in \mathfrak{k}$ ) to

$$\theta(X) = \theta(X_1 + iX_2) = X_1 - iX_2. \quad (2.4.52)$$

In this way,  $\mathfrak{k}$  can be identified with the +1 eigenspace of  $\theta$ , while  $i\mathfrak{k}$  is identified with its  $-1$  eigenspace. In addition, we can define a *local \*-involution*,  $*$  :  $\mathfrak{g} \rightarrow \mathfrak{g}$ , on  $\mathfrak{g}$  as the map sending  $X \in \mathfrak{g}$  to

$$X^* = -\theta(X) = -X_1 + iX_2. \quad (2.4.53)$$

In this way, if  $\mathfrak{k} = \mathfrak{u}(d)$  is the Lie algebra of  $d \times d$  skew-Hermitian matrices, then  $i\mathfrak{k} = i\mathfrak{u}(d)$  is the Lie algebra of  $d \times d$  Hermitian matrices (with negated Lie bracket), and then the  $*$ -involution coincides with the complex conjugate of a complex matrix in the sense that

$$\begin{aligned} X \in \mathfrak{k} &\iff X^* = -\theta(X) = -X, \\ X \in i\mathfrak{k} &\iff X^* = -\theta(X) = +X. \end{aligned} \quad (2.4.54)$$

Finally, it is sometimes useful to give the factor  $i\mathfrak{k}$  its own dedicated symbol, namely

$$\mathfrak{p} := i\mathfrak{k}, \quad (2.4.55)$$

and thus we have  $\mathfrak{g} = \mathfrak{k}_{\mathbb{C}} = \mathfrak{k} \oplus i\mathfrak{k} = \mathfrak{k} \oplus \mathfrak{p}$ .

**Remark 2.4.29.** In addition to the notational conventions from Remark 2.4.28 for dealing with the complexification,  $\mathfrak{g} = \mathfrak{k}_{\mathbb{C}} = \mathfrak{k} \oplus i\mathfrak{k}$ , of a real Lie algebra  $\mathfrak{k}$ , it will also be important, for the sake of consistency, to establish a notational convention for dealing with the subalgebras of  $\mathfrak{k}$  and their complexification. In particular, the maximal abelian subalgebra (or Cartan subalgebra) of  $\mathfrak{k}$ , will always be denoted by  $\mathfrak{t}$ , in which case we have the relationship

$$\mathfrak{t} \subseteq \mathfrak{k}. \quad (2.4.56)$$

The complexification of  $\mathfrak{t}$ , will be denoted by

$$\mathfrak{h} = \mathfrak{t}_{\mathbb{C}} = \mathfrak{t} \oplus i\mathfrak{t}. \quad (2.4.57)$$

Evidently, just as  $\mathfrak{t}$  is a subalgebra of  $\mathfrak{k}$ , the complexification of  $\mathfrak{t}$  is a subalgebra of the complexification of  $\mathfrak{k}$ ,

$$\mathfrak{h} = \mathfrak{t} \oplus i\mathfrak{t} \subseteq \mathfrak{k} \oplus i\mathfrak{k} = \mathfrak{g}. \quad (2.4.58)$$

Furthermore,  $\mathfrak{h}$  is the maximal abelian subalgebra of  $\mathfrak{g}$  [Hal15, Prop. 7.11]. Finally, in addition to sometimes letting  $\mathfrak{p} := i\mathfrak{k}$ , we sometimes let  $\mathfrak{a} := i\mathfrak{t}$ , in which case  $\mathfrak{a} \subseteq \mathfrak{p}$ .

**Example 2.4.30.** An key example to illustrate all of these notational conventions is to start with the Lie algebra  $\mathfrak{k} = \mathfrak{u}(d)$ . In this case, we have

- $\mathfrak{k} = \mathfrak{u}(d)$  (skew-Hermitian  $d \times d$  matrices),
- $\mathfrak{p} = i\mathfrak{k} = i\mathfrak{u}(d)$  (Hermitian  $d \times d$  matrices),
- $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p} = \mathfrak{gl}(d, \mathbb{C})$  ( $d \times d$  complex matrices),
- $\mathfrak{t} = \mathfrak{u}(1)^{\times d} \cong i\mathbb{R}^d$  ( $d \times d$  diagonal matrices with imaginary entries),
- $\mathfrak{a} = i\mathfrak{t} = i\mathfrak{u}(1)^{\times d} \cong \mathbb{R}^d$  ( $d \times d$  diagonal matrices with real entries), and
- $\mathfrak{h} = \mathfrak{t} \oplus \mathfrak{a} \cong \mathbb{C}^d$  ( $d \times d$  diagonal matrices with complex entries).

**Definition 2.4.31.** A Lie algebra,  $\mathfrak{g}$ , is called *reductive* if

$$\mathfrak{g} = \mathfrak{k}_{\mathbb{C}} \quad (2.4.59)$$

where  $\mathfrak{k}$  is the Lie algebra of a compact matrix Lie group  $K$ . If  $\mathfrak{g}$  is reductive and has trivial center, i.e.,  $\mathfrak{z}(\mathfrak{g}) = 0$ , then it is *semisimple*.

**Example 2.4.32.** Note that every complex reductive Lie algebra  $\mathfrak{g}$  can be decomposed as a direct sum

$$\mathfrak{g} = \mathfrak{z}(\mathfrak{g}) \oplus \mathfrak{s}, \quad (2.4.60)$$

where  $\mathfrak{z}(\mathfrak{g})$  is its center and  $\mathfrak{s}$  is semisimple [Hal15, Prop. 7.6]. For example, the Lie algebra of  $d \times d$  complex matrices,  $\mathfrak{gl}(d, \mathbb{C})$ , is reductive as it is the complexification of  $\mathfrak{gl}(d, \mathbb{R})$  or  $\mathfrak{u}(d)$ , but it is not semisimple because its center,

$$\mathfrak{z}(\mathfrak{gl}(d, \mathbb{C})) = \{z\mathbb{1}_d \mid z \in \mathbb{C}\} \cong \mathbb{C}, \quad (2.4.61)$$

consists of the set of scalar  $d \times d$  matrices. In this case, Eq. (2.4.60) becomes

$$\mathfrak{gl}(d, \mathbb{R}) \cong \mathbb{C} \oplus \mathfrak{sl}(d, \mathbb{C}), \quad (2.4.62)$$

where semisimple factor,  $\mathfrak{sl}(d, \mathbb{C})$ , consists of all traceless  $d \times d$  complex matrices.

Although finite-dimensional representations of semisimple Lie algebras can be shown to be completely reducible [Hal15, Thm. 10.9], the same is not true of finite-dimensional representations of reductive Lie algebras [Hal15, Ex. 10.9.1]<sup>2</sup>. Given that finite-dimensional

---

<sup>2</sup>A standard example is the Lie algebra representation  $\phi : \mathbb{C} \rightarrow \mathfrak{gl}(2, \mathbb{C})$  sending  $z \in \mathbb{C}$  to  $\phi(z) = \begin{pmatrix} 0 & z \\ 0 & 0 \end{pmatrix}$  which contains non-trivial invariant subspaces but is not expressible as the direct sum of smaller representations of  $\mathbb{C}$ .

representations of semisimple Lie algebras are completely reducible [Hal15, Thm. 10.9], it becomes worth classifying their finite-dimensional irreducible representations. This classification is provided by the theorem of highest weights (Theorem 2.4.49) and depends on a handful of supporting concepts and terminology to be covered over the next few pages.

The crucial reason for taking  $\mathfrak{k}$  in the definition of a reductive Lie algebra to be the Lie algebra of a compact Lie group  $K$  is to import Weyl's unitary trick mentioned in the proof of Theorem 2.4.25 for the purposes of establishing the following result [Hal15, Prop. 7.4].

**Lemma 2.4.33.** *Let  $\mathfrak{g} = \mathfrak{k}_{\mathbb{C}}$  be a reductive Lie algebra. Then there exists an inner product  $\langle \cdot, \cdot \rangle : \mathfrak{g} \times \mathfrak{g} \rightarrow \mathbb{C}$  on  $\mathfrak{g}$  which is (i) real-valued when restricted to  $\mathfrak{k} \subseteq \mathfrak{g}$ , and (ii) satisfies for all  $X, Y, Z \in \mathfrak{g}$ ,*

$$\langle \text{ad}(X)(Y), Z \rangle = \langle Y, \text{ad}(X^*)(Z) \rangle, \quad (2.4.63)$$

where  $\text{ad} : \mathfrak{g} \rightarrow \mathfrak{gl}(\mathfrak{g})$  is the adjoint representation of  $\mathfrak{g}$  on itself, and  $X^*$  is defined as in Eq. (2.4.53).

Recall, at this stage, that the roots of a Lie algebra are the weights (simultaneous eigenvalues) of its adjoint representation (recall Definition 2.4.14 and Example 2.4.15). In the context of reductive Lie algebras, roots must interact with the complex structure in a particular way.

**Lemma 2.4.34.** *Let  $\alpha \in \mathfrak{h}^*$  be a root for a reductive Lie algebra  $\mathfrak{g} = \mathfrak{k} \oplus i\mathfrak{k}$  relative to a maximal abelian subalgebra  $\mathfrak{h} = \mathfrak{t} \oplus i\mathfrak{t}$ . If  $H = H_1 + iH_2 \in \mathfrak{h}$  for  $H_1, H_2 \in \mathfrak{t}$ , then*

$$\alpha(H)^* = -\alpha(H_1) + i\alpha(H_2). \quad (2.4.64)$$

*Proof.* Let  $X \in \mathfrak{g}$  be a non-zero eigenvector of the adjoint representation of  $H \in \mathfrak{t} \subseteq \mathfrak{k}$  with eigenvalue  $\alpha \in \mathbb{C}$ , meaning

$$\text{ad}(H)(X) = [H, X] = \alpha X. \quad (2.4.65)$$

Since the inner product in Lemma 2.4.33 is anti-linear in the first argument and linear in the second argument, and  $H \in \mathfrak{t} \subseteq \mathfrak{k}$  satisfies  $H^* = -H$ , one concludes

$$\begin{aligned} \alpha^* \langle X, X \rangle &= \langle \alpha X, X \rangle, \\ &= \langle \text{ad}(H)(X), X \rangle, \\ &= \langle X, \text{ad}(H^*)(X) \rangle, \\ &= -\langle X, \text{ad}(H)(X) \rangle, \\ &= -\alpha \langle X, X \rangle. \end{aligned} \quad (2.4.66)$$

Therefore, for every  $H \in \mathfrak{t}$ , its adjoint representation,  $\text{ad}(H)$ , has purely imaginary eigenvalues,  $\alpha^* = -\alpha$ . Since roots are just simultaneous eigenvectors of the adjoint representation of  $\mathfrak{h}$ , one concludes, for all  $H \in \mathfrak{t}$ , that  $\alpha(H) \in i\mathbb{R}$  is imaginary. Therefore,

$$\alpha(H)^* = \alpha(H_1 + iH_2)^* = (\alpha(H_1) + i\alpha(H_2))^*, \quad (2.4.67)$$

$$= \alpha(H_1)^* - i\alpha(H_2)^* = -\alpha(H_1) + i\alpha(H_2). \quad (2.4.68)$$

□

**Remark 2.4.35.** In light of [Lemma 2.4.34](#), a root of a reductive Lie algebra, while having the type of a complex-linear function  $\alpha : \mathfrak{h} \rightarrow \mathbb{C}$  (i.e.  $\alpha \in \mathfrak{h}^*$ ), is essentially determined by its restriction to  $\mathfrak{t}$  or  $i\mathfrak{t}$ . Specifically, every root can be identified with a real-linear function on  $i\mathfrak{t}$ , i.e.  $\mathfrak{h} : i\mathfrak{t} \rightarrow \mathbb{R}$ .

Understanding the algebraic structure of the roots and associated root spaces yields the most significant result in the representation theory of semisimple Lie algebras.

**Theorem 2.4.36.** *Let  $\mathfrak{g} = \mathfrak{k}_{\mathbb{C}}$  be a complex semisimple Lie algebra with fixed maximal abelian subalgebra  $\mathfrak{h}$ . Let  $R \subseteq \mathfrak{h}^*$  denote the set of all non-zero roots of  $\mathfrak{g}$ . Then  $\mathfrak{g}$  can be decomposed as the direct sum*

$$\mathfrak{g} = \mathfrak{h} \oplus \bigoplus_{\alpha \in R} \mathfrak{g}_{\alpha}, \quad (2.4.69)$$

where  $\mathfrak{g}_{\alpha} \subseteq \mathfrak{g}$  is the root space with associated to  $\alpha \in \mathfrak{h}^*$ .

Furthermore,

- for all  $\alpha, \beta \in \mathfrak{h}^*$ ,  $[g_{\alpha}, g_{\beta}] \subseteq g_{\alpha+\beta}$ , meaning

$$X \in g_{\alpha}, Y \in g_{\beta} \implies [X, Y] \in g_{\alpha+\beta}, \quad (2.4.70)$$

- if  $\alpha \in R$  is a root with root vector  $X \in g_{\alpha}$ , then  $-\alpha \in \mathfrak{h}^*$  is also a root with root vector  $X^* \in g_{-\alpha}$ , and moreover  $\alpha$  and  $-\alpha$  are the only non-zero roots proportional to  $\alpha$ ,
- the roots span all of  $\mathfrak{h}^*$ ,

$$\text{span}(R) = \mathfrak{h}^*, \quad (2.4.71)$$

- for each root  $\alpha \in R$ , the root space,  $\mathfrak{g}_{\alpha}$ , is one-dimensional,

$$\dim(\mathfrak{g}_{\alpha}) = 1. \quad (2.4.72)$$

*Proof.* The statement of the above theorem is a combination of a number of results [[Hal15](#), Prop. 7.16-7.18, Thm. 7.19, Thm 7.23].  $\square$

**Theorem 2.4.37.** *Let  $\mathfrak{g} = \mathfrak{k}_{\mathbb{C}}$  be a reductive Lie algebra. Then there exists a subalgebra  $\mathfrak{s}^{\alpha} \subseteq \mathfrak{g}$  spanned by  $H_{\alpha} \in \mathfrak{h}$ ,  $X_{\alpha} \in \mathfrak{g}_{\alpha}$ , and  $Y_{\alpha} = X_{\alpha}^* \in \mathfrak{g}_{-\alpha}$  which satisfies*

$$[H_{\alpha}, X_{\alpha}] = 2X_{\alpha}, \quad [H_{\alpha}, Y_{\alpha}] = 2Y_{\alpha}, \quad [X_{\alpha}, Y_{\alpha}] = H_{\alpha}, \quad (2.4.73)$$

and thus  $\mathfrak{s}^{\alpha}$  is isomorphic to  $\mathfrak{sl}(2, \mathbb{C})$  from [Example 2.3.19](#). Moreover,  $H_{\alpha} \in \mathfrak{h}$  is uniquely determined by  $\alpha$  and called the **coroot** of  $\alpha$ .

*Proof.* Let  $\langle \cdot, \cdot \rangle$  be the inner product on  $\mathfrak{g}$  provided by [Lemma 2.4.33](#) Then let  $X \in g_{\alpha}$  be a non-zero root vector for the root  $\alpha \in \mathfrak{h}^*$ , meaning

$$\forall H \in \mathfrak{h} : \text{ad}(H)(X) = [H, X] = \alpha(H)X. \quad (2.4.74)$$

By [Theorem 2.4.36](#),  $X^* \in g_{-\alpha}$  is a root vector for the root  $-\alpha \in \mathfrak{h}^*$ . Therefore, by [Lemma 2.4.33](#), for all  $H \in \mathfrak{h}$ ,

$$\langle [X, X^*], H \rangle = \langle X^*, [X^*, H] \rangle = -\langle X^*, [H, X^*] \rangle = \alpha(H) \langle X^*, X^* \rangle. \quad (2.4.75)$$



As  $X^*$  is non-zero, so is  $\langle X^*, X^* \rangle$ , which means  $\alpha$  is proportional to  $[X, X^*]$  in the sense that

$$\forall H \in \mathfrak{h} : \alpha(H) = \frac{\langle [X, X^*], H \rangle}{\langle X^*, X^* \rangle}. \quad (2.4.76)$$

Again using [Theorem 2.4.36](#),  $[X, X^*] \in g_0 = \mathfrak{h}$  and therefore

$$\alpha([X, X^*]) = \frac{\|[X, X^*]\|^2}{\|X^*\|^2} > 0 \quad (2.4.77)$$

is well-defined, real and positive. The first claim of the proof follows from defining

$$H_\alpha = \frac{2}{\alpha([X, X^*])} [X, X^*], \quad X_\alpha = \sqrt{\frac{2}{\alpha([X, X^*])}} X, \quad Y_\alpha = \sqrt{\frac{2}{\alpha([X, X^*])}} X^*, \quad (2.4.78)$$

which satisfy the claimed commutation relations. To prove that  $H_\alpha$  is independent of the choice of non-zero root vector  $X \in g_\alpha$  initially chosen, it suffices to note that, by [Theorem 2.4.36](#),  $\dim(g_\alpha) = 1$ , and thus the only other choice would be a scalar multiple of  $X$  which leaves the definition of  $H_\alpha$  unaffected as  $\alpha$  is proportional to  $H_\alpha$ ,

$$\alpha(H) \propto \langle H_\alpha, H \rangle, \quad (2.4.79)$$

with proportionality constant  $\|[X, X^*]\|^2 / (2 \|X^*\|^4)$  which is independent of scalar multiples of  $X$ .  $\square$

**Corollary 2.4.38.** *Let  $\mathfrak{g} = \mathfrak{k}_\mathbb{C}$  be a finite-dimensional reductive Lie algebra, let  $\alpha, \beta \in \mathfrak{h}^*$  be roots of  $\mathfrak{g}$  and let  $H_\beta \in \mathfrak{h}$  be the coroot of  $\beta$  as defined in [Theorem 2.4.37](#). Then  $\alpha(H_\beta)$  is an integer,*

$$\alpha(H_\beta) \in \mathbb{Z}. \quad (2.4.80)$$

*Proof.* Given that the coroot  $H_\beta \in \mathfrak{h}$  as defined by [Theorem 2.4.37](#), belongs to subalgebra of  $\mathfrak{g}$  isomorphic to  $\mathfrak{sl}(2, \mathbb{C})$ , we can apply [Example 2.4.13](#) to see that the eigenvalues of  $\phi(H_\beta)$  for any finite-dimensional representation must be integers. Finally, given that  $\mathfrak{g}$  is assumed finite-dimensional, and that  $\alpha(H_\beta)$  is an eigenvalue of  $\text{ad}(H_\beta)$ , it must be that  $\alpha(H_\beta)$  is an integer.  $\square$

**Remark 2.4.39.** The results of [Theorem 2.4.36](#), [Corollary 2.4.38](#) and [[Hal15](#), Thm. 7.26] imply that the roots of any finite-dimensional semisimple complex Lie algebra constitute an abstract *root system* (see [Section 2.4.6](#)). Using the inner product provided by [Lemma 2.4.33](#), it is possible to identify  $\mathfrak{h}$  with its dual vector space  $\mathfrak{h}^*$  and therefore every root  $\alpha \in \mathfrak{h}^*$  with an unique element  $H'_\alpha \in \mathfrak{h}$  such that

$$\alpha(H) = \langle H'_\alpha, H \rangle, \quad (2.4.81)$$

holds for every  $H \in \mathfrak{h}'$ . By [Theorem 2.4.37](#) we see that the element  $H'_\alpha \in H$  would be *proportional* to the coroot  $H_\alpha \in H$ . The convention adopted in this thesis is to consider roots as functionals, i.e.  $\alpha \in \mathfrak{h}^*$ , in which case the enveloping vector space for the abstract root systems considered in [Section 2.4.6](#) should be taken to be  $\mathcal{E} = \mathfrak{h}^*$  with inner product,  $\langle \cdot, \cdot \rangle'$  on  $\mathfrak{h}'$ , *dual* to the inner product provided by [Lemma 2.4.33](#), i.e., for all  $\alpha, \beta \in \mathfrak{h}^*$  let

$$\langle \alpha, \beta \rangle' := \langle H'_\alpha, H'_\beta \rangle. \quad (2.4.82)$$

Of course, in the setting of finite-dimensional inner product spaces, which convention one chooses is purely a matter of preference.



## 2.4.6 Roots systems

**Definition 2.4.40.** A **root system**  $R$  is a finite set of non-zero vectors, called **roots**, of a finite dimensional real inner product space  $(\mathcal{E}, \langle \cdot, \cdot \rangle)$  such that

- i) the roots span  $\mathcal{E}$ , i.e.  $\text{span}(R) = \mathcal{E}$ ,
- ii) if  $\alpha \in R$ , then  $s\alpha \in R$  if and only if  $s \in \{-1, +1\}$ ,
- iii) if  $\alpha, \beta \in R$ , then  $c_\alpha(\beta)$  defined by

$$c_\alpha(\beta) := 2 \frac{\langle \alpha, \beta \rangle}{\langle \alpha, \alpha \rangle}, \quad (2.4.83)$$

in an integer (called a **Cartan integer**), and

- iv) if  $\alpha, \beta \in R$ , then  $r_\alpha(\beta) \in R$  where

$$r_\alpha(\beta) := \beta - c_\alpha(\beta)\alpha. \quad (2.4.84)$$

The **rank of a root system**  $R$  is the dimension of  $\mathcal{E}$ .

**Remark 2.4.41.** For any root  $\alpha \in R$  of a root system  $R$ , Eq. (2.4.84) defines a *reflection*  $r_\alpha : R \rightarrow R$  sending each root  $\beta \in R$  to the root  $r_\alpha(\beta) \in R$  obtained by reflecting  $\beta$  through the hyperplane  $\mathcal{E}_\alpha \subset \mathcal{E}$  (of dimension  $\dim(\mathcal{E}) - 1$ ) orthogonal to  $\alpha$  defined by

$$\mathcal{E}_\alpha := \{v \in \mathcal{E} \mid \langle \alpha, v \rangle = 0\}. \quad (2.4.85)$$

The collection of all such reflections  $W = \{r_\alpha \mid \alpha \in R\}$  generates a finite subgroup of the orthogonal group on  $\mathcal{E}$  known as the **Weyl group** of the root system  $R$ .

**Definition 2.4.42.** Let  $R \subset \mathcal{E}$  be a root system of the real inner product space  $(\mathcal{E}, \langle \cdot, \cdot \rangle)$  and let  $S : \mathcal{E} \rightarrow \{-1, 0, +1\}^R$  be the function assigning to each element  $\mu \in \mathcal{E}$  the function,  $S_\mu : R \rightarrow \{-1, 0, +1\}$ , mapping each root  $\alpha \in R$  to the sign of  $\langle \mu, \alpha \rangle$ :

$$S_\mu(\alpha) = \text{sign}(\langle \mu, \alpha \rangle) = \begin{cases} -1 & \text{if } \langle \mu, \alpha \rangle < 0, \\ 0 & \text{if } \langle \mu, \alpha \rangle = 0, \\ +1 & \text{if } \langle \mu, \alpha \rangle > 0. \end{cases} \quad (2.4.86)$$

Since there are only finitely many  $\{-1, 0, +1\}$ -valued functions on  $R$ , the map  $\mu \mapsto S_\mu$  partitions  $\mathcal{E}$  into finitely many disjoint regions. Those regions where  $S_\mu$  is strictly non-zero for all roots,

$$\forall \alpha \in R : S_\mu(\alpha) = \text{sign}(\langle \mu, \alpha \rangle) \neq 0, \quad (2.4.87)$$

are called the **Weyl chambers**. In other words, the Weyl chambers are the open connected components of the set-wise difference  $\mathcal{E} \setminus \bigcup_{\alpha \in R} \mathcal{E}_\alpha$  where  $\mathcal{E}_\alpha$  is the hyperplane orthogonal to the root  $\alpha$  as defined by Eq. (2.4.85).

The Weyl group,  $W$ , acts transitively on the Weyl chambers [Hal15, Prop. 8.23].

The following remark demonstrates that for every root system  $R$  one can always pick a subset of roots  $\Delta \subset R$  called a *base* for  $R$  that has a number of desirable properties [Hal15, Thm. 8.16]. Choosing a base for a root system, as we shall see, is equivalent to choosing a particular Weyl chamber to be the *positive* one.

**Remark 2.4.43.** Let  $R \subset \mathcal{E}$  be a root system in  $\mathcal{E}$  and fix a hyperplane  $\mathcal{W} \subset \mathcal{E}$  of dimension  $\dim(\mathcal{E}) - 1$ , called the *separating hyperplane*, such that no root  $\alpha \in R$  is contained within  $\mathcal{W}$ . That such a hyperplane actually exists follows from choosing a Weyl chamber, denoted by  $C$ , to be considered as the *positive Weyl chamber*, and then choosing a vector  $\omega \in C$  belonging to the positive Weyl chamber to serve as the normal vector of  $\mathcal{W}$ , i.e.  $\mathcal{W}$  is defined by  $\omega \in C$  by

$$\mathcal{W} = \{v \in \mathcal{E} \mid \langle \omega, v \rangle = 0\}. \quad (2.4.88)$$

As  $\omega$  belongs to the interior of a Weyl chamber, it is not contained in any of the hyperplanes  $\mathcal{E}_\alpha$  defined by Eq. (2.4.85) [Hal15, Prop. 8.14]. The separating hyperplane  $\mathcal{W}$  derives its name because every root  $\alpha \in R$  belongs exactly one of the two connected components in  $\mathcal{E} \setminus \mathcal{W}$ . In fact, given a hyperplane  $\mathcal{W}$  with normal  $\omega$ , a root  $\alpha \in R$  is said to be a *positive root* if  $\langle \alpha, \omega \rangle > 0$  or a *negative root* if  $\langle \alpha, \omega \rangle < 0$ . The set of positive roots in  $R$  is denoted  $R_+$  and the set of negative roots is denoted  $R_-$  such that  $R$  is the disjoint union  $R = R_+ \sqcup R_-$ . While this bipartition of the root system into positive and negative roots evidently depends on the initial choice of positive Weyl chamber, it is independent of the choice of separating hyperplane.

With respect to this decomposition of the root system, the subset  $\Delta \subset R_+ \subset R$  of positive roots that cannot be written as a sum of two or more positive roots is called the *base* of the root system. Using this construction, one can additionally show that i) the elements of the base  $\Delta$  are linearly independent and ii) every positive (respectively negative) root is expressible as a non-negative (respectively non-positive) integer linear combination of elements in  $\Delta$  [Hal15, Thms. 8.16]. These final two conditions are sometimes taken as an axiomatic definition for a *base of a root system*, but it can be shown that every subset  $\Delta \subset R$  satisfying these conditions arises under the construction presented here for some (non-unique) choice of separating hyperplane  $\mathcal{W}$  and normal  $\omega$  [Hal15, Thms. 8.16 & 8.17].

The next two definitions are concerned with elements  $\mu \in \mathcal{E}$  of the enveloping vector space that are not necessarily roots.

**Definition 2.4.44.** Let  $R \subset \mathcal{E}$  be a root system with inner product  $\langle \cdot, \cdot \rangle$ . For each root,  $\alpha \in R$ , its *coroot* is defined by

$$H_\alpha := 2 \frac{\alpha}{\langle \alpha, \alpha \rangle}. \quad (2.4.89)$$

An element  $\mu \in \mathcal{E}$  is called an *algebraically integral element* if for all roots  $\alpha \in R$ ,

$$\langle \mu, H_\alpha \rangle = 2 \frac{\langle \mu, \alpha \rangle}{\langle \alpha, \alpha \rangle} \in \mathbb{Z}. \quad (2.4.90)$$

**Definition 2.4.45.** Let  $R \subset \mathcal{E}$  be a root system with inner product  $\langle \cdot, \cdot \rangle$  and let  $\Delta \subset R$  be a base for  $R$  (Remark 2.4.43). An element  $\mu \in \mathcal{E}$  is called *dominant* if

$$\forall \alpha \in \Delta : \langle \mu, \alpha \rangle \geq 0, \quad (2.4.91)$$

and *strictly dominant* if,

$$\forall \alpha \in \Delta : \langle \mu, \alpha \rangle > 0. \quad (2.4.92)$$

**Remark 2.4.46.** Relative to a given choice of base,  $\Delta$ , for a root system,  $R \subset \mathcal{E}$ , it is not too difficult to see that an element  $\mu \in \mathcal{E}$  is strictly dominant if and only if it belongs to the positive Weyl chamber (Remark 2.4.43) and dominant if and only if it belongs to the closure of the positive Weyl chamber.

**Definition 2.4.47.** Let  $R \subset \mathcal{E}$  be a root system with bases  $\Delta = \{\alpha_1, \dots, \alpha_r\} \subset R$ . An element  $\mu \in \mathcal{E}$  is *higher* than an element  $\nu \in \mathcal{E}$ , expressed as

$$\mu \geq \nu, \quad (2.4.93)$$

if there exists non-negative integers  $\{c_1, \dots, c_r\}$  such that

$$\mu - \nu = c_1\alpha_1 + \dots + c_r\alpha_r. \quad (2.4.94)$$

The relation,  $\geq$ , defines a partial ordering on  $\mathcal{E}$ .

**Example 2.4.48.** Let  $n \in \mathbb{N}$  and let  $\{e_1, e_2, \dots, e_n, e_{n+1}\}$  be the standard orthonormal basis for  $\mathbb{R}^{n+1}$  and let  $\mathcal{E} \simeq \mathbb{R}^n$  be the subspace of vectors  $v$  with components  $v_j = \langle v, e_j \rangle$  summing to zero. Then the  $A_n$  root system is the root system consisting of all vectors of the form  $\alpha_{ij} = e_i - e_j$  for some  $i \neq j \in [n+1]$ . The Cartan integer is given by

$$c_{\alpha_{ij}}(\alpha_{kl}) = 2 \frac{\langle \alpha_{ij}, \alpha_{kl} \rangle}{\langle \alpha_{ij}, \alpha_{ij} \rangle} = \langle \alpha_{ij}, \alpha_{kl} \rangle = \delta_{ik} + \delta_{jl} - \delta_{il} - \delta_{jk}. \quad (2.4.95)$$

The base  $\Delta$  for the  $A_n$  root system is typically taken to be the of roots  $\alpha_{ij}$  with  $j = i + 1$ :

$$\Delta = \{\alpha_{12}, \alpha_{23}, \dots, \alpha_{n-1,n}, \alpha_{n,n+1}\}. \quad (2.4.96)$$

With respect to this standard base for  $A_n$ , the positive roots are those roots  $\alpha_{ij}$  with  $i > j$  and the negative roots are those with  $i < j$ . The Weyl group  $W$  for the  $A_n$  is given by the faithful representation of the symmetric group  $S_n$  which acts on the roots,  $\alpha_{ij}$ , for  $\pi \in S_n$  by  $\alpha_{ij} \mapsto \alpha_{\pi(i)\pi(j)}$ . This representation of  $S_n$  is also known as the *standard representation* of  $S_n$ . The root system  $A_n$  is also the root system associated to the Lie algebra  $\mathfrak{sl}(n+1, \mathbb{C})$ .

## 2.4.7 Highest weights

The finite-dimensional irreducible representations of complex semisimple Lie algebras are determined, up to isomorphism, by their highest weights. The following theorem is one part of the theorem of highest weights for complex semisimple Lie algebras.

**Theorem 2.4.49.** *Let  $\mathfrak{g}$  be a complex semisimple Lie algebra, let  $\mathfrak{h} \subseteq \mathfrak{g}$  be a fixed maximal abelian subalgebra and let  $R$  be the root system of  $\mathfrak{g}$  (relative to  $\mathfrak{h}$ ), and let  $\Delta \subset R$  be a base for  $R$  with positive roots  $R_+$  and negative roots  $R_-$ . If  $\phi : \mathfrak{g} \rightarrow \mathfrak{gl}(\mathcal{V})$  is an irreducible representation of  $\mathfrak{g}$  on a finite-dimensional complex vector space  $\mathcal{V}$ , then exists a weight  $\lambda \in \mathfrak{h}^*$  for  $\phi$  with weight vector  $v_\lambda \in \mathcal{V}_\lambda \subset \mathcal{V}$ , meaning*

$$\forall H \in \mathfrak{h} : \phi(H)v_\lambda = \lambda(H)v_\lambda, \quad (2.4.97)$$

such that  $\lambda$  is a **highest weight**, meaning for all positive roots  $\alpha \in R_+$  and corresponding root vectors  $X \in \mathfrak{g}_\alpha$ ,

$$\phi(X)v_\lambda = 0. \quad (2.4.98)$$

Moreover, the highest weight  $\lambda \in \mathfrak{h}^*$  is a dominant, algebraically integral element relative to the root system  $R$  and base  $\Delta$ . Furthermore, if  $\phi' : \mathfrak{g} \rightarrow \mathfrak{gl}(\mathcal{V}')$  is another finite-dimensional irreducible representation of  $\mathfrak{g}$  with highest weight  $\lambda$ , then  $\phi'$  and  $\phi$  are isomorphic representations of  $\mathfrak{g}$ .

*Proof.* Let  $\lambda \in \mathfrak{h}^*$  be any weight for  $\phi$  with weight vector  $v \in \mathcal{V}$ . If  $w = \phi(X)v \neq 0$  was non-zero for some root vector  $X \in \mathfrak{g}_\alpha$  with root  $\alpha \in R$ , it would constitute a weight vector with weight  $\alpha + \lambda \in \mathfrak{h}^*$  because for all  $H \in \mathfrak{h}$ ,

$$\phi(H)\phi(X)v = \phi([H, X])v + \phi(X)\phi(H)v = \alpha(H)\phi(X)v - \lambda(H)\phi(X)v, \quad (2.4.99)$$

and thus

$$\phi(H)w = (\alpha + \lambda)(H)w. \quad (2.4.100)$$

By the finite-dimensionality of  $\mathcal{V}$ , can only be finitely many weights  $\lambda \in \mathfrak{h}^*$  for the representation  $\phi$  and therefore there must exist a weight that is highest in the sense that for all positive roots  $\alpha \in R_+$ , Eq. (2.4.98) holds.

That  $\lambda \in \mathfrak{h}^*$  is a dominant, algebraically integral element with respect to the root system  $R$  with base  $\Delta$  follows from first recalling Corollary 2.4.38 (which guarantees  $\lambda$  is an algebraically integral element) and then second noting that if  $\lambda$  is not dominant, e.g.  $\lambda(H_\alpha) < 0$  for some root  $\alpha \in \Delta$  in the base  $\Delta$ , then  $\phi(X)v \neq 0$  is necessarily a weight vector with weight  $\lambda + \alpha$ , contradicting Eq. (2.4.98).

That irreducible representations with the same highest weight are isomorphic follows from noting that if  $v_\lambda \in \mathcal{V}$  and  $v'_\lambda \in \mathcal{V}'$  are highest weight vectors with the same weight  $\lambda \in \mathfrak{h}^*$ , then  $v_\lambda \oplus v'_\lambda$  is a highest weight vector with weight  $\lambda$  for the direct sum representation  $\phi \oplus \phi'$  of  $\mathfrak{g}$  on  $\mathcal{V} \oplus \mathcal{V}'$ . By restricting  $\phi \oplus \phi'$  to the invariant subspace  $\mathcal{U} \subseteq \mathcal{V} \oplus \mathcal{V}'$  containing  $v_\lambda \oplus v'_\lambda$  and applying a version of Schur's lemma for Lie algebras, we see that both  $\mathcal{V}$  and  $\mathcal{V}'$  are isomorphic to the restriction of  $\phi \oplus \phi'$  onto  $\mathcal{U}$  and thus are isomorphic to each other [Hal15, Prop 6.15].  $\square$

**Remark 2.4.50.** Although weights,  $\lambda : \mathfrak{h} \rightarrow \mathbb{C}$ , assign complex numbers,  $\lambda(H)$ , to elements  $H \in \mathfrak{h}$ , of the complexified Lie algebra  $\mathfrak{h} = \mathfrak{t} \oplus i\mathfrak{t}$ , they are uniquely determined by their assignment of purely imaginary numbers to  $\mathfrak{t}$ , or equivalently, by their assignment of real numbers to  $i\mathfrak{t}$ , i.e.

$$H \in \mathfrak{t} \implies \lambda(H) \in i\mathbb{R}, \quad (2.4.101)$$

$$H \in i\mathfrak{t} \implies \lambda(H) \in \mathbb{R}. \quad (2.4.102)$$

This observation is analogous to the result of Lemma 2.4.34 for roots. In this manner, weights can be identified by their restriction to either  $\mathfrak{t}$  or  $i\mathfrak{t}$ . In the latter case, where  $\lambda$  is identified with its restriction to  $i\mathfrak{t}$ , denoted by  $\lambda \in i\mathfrak{t} \rightarrow \mathbb{R}$ , the weight  $\lambda$  is called a *real weight* and written as  $\lambda \in (i\mathfrak{t})^*$ .

It can also be shown that a converse to [Theorem 2.4.49](#) also holds and states that for every dominant algebraically integral element  $\lambda \in \mathfrak{h}^*$  is the highest weight of some finite-dimensional irreducible representation [[Hal15](#), Thm. 9.5]. [Theorem 2.4.49](#) and its converse, when taken together, states that the isomorphism classes of finite-dimensional irreducible representations of complex semisimple Lie algebras are in bijection with the set of dominant algebraically integral elements  $\lambda \in \mathfrak{h}^*$ .

Not only does the theorem of highest weights provide a complete classification of the finite-dimensional irreducible representations for semisimple Lie algebras, it can be extended to provide a complete classification of the finite-dimensional irreducible representations of compact connected Lie groups and their complexifications. The reason for this correspondence lies with the observation that if  $K$  is a compact connected Lie group with Lie algebra  $\mathfrak{k}$ , the exponential map  $\exp : \mathfrak{k} \rightarrow K$ , which assigns to each element  $X \in \mathfrak{k}$  of the Lie algebra an element  $\exp(X) \in K$  of the Lie group, is *surjective* [[Hal15](#), Cor. 11.10], i.e.

$$K = \exp(\mathfrak{k}) := \{\exp(X) \in K \mid X \in \mathfrak{k}\}. \quad (2.4.103)$$

If  $K$  is not compact<sup>3</sup> or connected, then this is no longer true.

To begin, let  $T$  to be an abelian Lie group and  $\Phi : T \rightarrow \text{GL}(\mathcal{V})$  a representation of  $T$  on a finite-dimensional complex vector space  $\mathcal{V}$ . As  $T$  is abelian,  $[\Phi(t), \Phi(t')] = 0$  for all  $t, t' \in T$ , and therefore Schur's lemma dictates that every irreducible representation of  $T$  is one-dimensional whose images are subgroups of the complex multiplicative group  $\mathbb{C}_\times$ , e.g.  $\Phi : T \rightarrow \mathbb{C}_\times$ . If the group  $T$  is both abelian and *compact*, then an irreducible representation  $\Phi$  must map  $T$  into the circle group  $U(1) \subset \mathbb{C}_\times$ . If  $\mathfrak{t}$  is the Lie algebra of  $T$ , then every irreducible representation  $\Phi : T \rightarrow U(1)$  of  $T$  satisfies for all  $A \in \mathfrak{t}$ ,

$$\Phi(\exp(A)) = e^{\lambda(A)} \quad (2.4.104)$$

where  $\lambda : \mathfrak{t} \rightarrow i\mathbb{R}$  is a purely imaginary linear function on  $\mathfrak{t}$ . Alternatively, if  $H \in i\mathfrak{t}$ , then the irreducible representation  $\Phi : T \rightarrow U(1)$  must be of the form

$$\Phi(\exp(iH)) = e^{i\lambda(H)} \quad (2.4.105)$$

where  $\lambda : i\mathfrak{t} \rightarrow \mathbb{R}$  is a real linear function on  $i\mathfrak{t}$ .

Of course, if  $H \in i\mathfrak{t}$  is such that  $\exp(iH)$  is the identity element in  $T$ , then  $\lambda : i\mathfrak{t} \rightarrow \mathbb{R}$  must map  $H$  to an integer multiple of  $2\pi$  in order to ensure that  $\Phi$  is indeed a representation.

Alternatively, if one considers those elements  $H' \in i\mathfrak{t}$  such that  $\exp(2\pi iH')$  is the identity in  $T$ , then for the same reason  $\lambda : i\mathfrak{t} \rightarrow \mathbb{R}$  must map  $H'$  to an integer  $\lambda(H') \in \mathbb{Z}$ .

**Definition 2.4.51.** Let  $T$  be a compact connected abelian Lie group, called a *torus*, with Lie algebra  $\mathfrak{t}$  and identity element  $e \in T$ . Let  $\Gamma \subset i\mathfrak{t}$  be defined by

$$\Gamma := \{H \in i\mathfrak{t} : \exp(2\pi iH) = e\}. \quad (2.4.106)$$

A linear function  $\lambda : i\mathfrak{t} \rightarrow \mathbb{R}$  is said to be *analytically integral* if for all  $H \in \Gamma \subset i\mathfrak{t}$ ,

$$\lambda(H) \in \mathbb{Z}. \quad (2.4.107)$$

---

<sup>3</sup>For a counterexample in the case of non-compact  $K$ , consider that  $\begin{pmatrix} -1 & 1 \\ 0 & -1 \end{pmatrix} \in \text{SL}(2, \mathbb{C})$  is not the exponential of any matrix  $X \in \mathfrak{sl}(2, \mathbb{C})$ .

In this manner, every irreducible representation of a torus  $T$  can be identified with an analytically integral element  $\lambda : \mathfrak{it} \rightarrow \mathbb{R}$ .

**Lemma 2.4.52.** *Let  $\Phi : T \rightarrow \mathrm{U}(1)$  be an irreducible representation of a torus  $T$ . Then there exists an analytically integral linear function  $\lambda : \mathfrak{it} \rightarrow \mathbb{R}$  satisfying for all  $H \in \mathfrak{it}$ ,*

$$\Phi(\exp(iH)) = e^{i\lambda(H)}. \quad (2.4.108)$$

Moreover,  $\lambda : \mathfrak{it} \rightarrow \mathbb{R}$  is a weight for the Lie algebra representation,  $\phi : \mathfrak{t} \rightarrow i\mathbb{R}$ , induced by  $\Phi : T \rightarrow \mathrm{U}(1)$  and  $i\lambda(H) = \phi(iH)$ .

**Remark 2.4.53.** Every torus  $T$  is isomorphic to the standard torus  $\mathrm{U}(1)^r$  for some  $r \in \mathbb{N}$  called the *rank* of  $T$ . The Lie algebra  $\mathfrak{t}$  of a torus  $T$  of rank  $r$  is isomorphic to  $\mathfrak{t} \cong i\mathbb{R}^r$ , and thus  $\mathfrak{it} \cong \mathbb{R}^r$ . As every element of  $\mathrm{U}(1)^r$  can be expressed as

$$(e^{i\theta_1}, e^{i\theta_2}, \dots, e^{i\theta_r}) \quad (2.4.109)$$

for some  $(\theta_1, \dots, \theta_r) \in [0, 2\pi)^r \subset \mathbb{R}^r = \mathfrak{it}$ , every irreducible representation of  $\mathrm{U}(1)^r$  is of the form

$$\Phi(e^{i\theta_1}, e^{i\theta_2}, \dots, e^{i\theta_r}) = e^{i(\lambda_1\theta_1 + m_2\theta_2 + \dots + m_r\theta_r)} \quad (2.4.110)$$

for some  $(m_1, \dots, m_r) \in \mathbb{Z}^r$ . In this way, the analytically integral weight,  $\lambda : \mathbb{R}^r \rightarrow \mathbb{R}$ , is the function

$$\lambda(\theta_1, \theta_2, \dots, \theta_r) = m_1\theta_1 + m_2\theta_2 + \dots + m_r\theta_r. \quad (2.4.111)$$

**Remark 2.4.54.** A torus  $T$  is called a **maximal torus** of a compact connected Lie group  $K$  if it is a subgroup  $T \subseteq K$  and if it is not a proper subgroup of any other torus in  $K$ . If  $T$  is a maximal torus of  $K$ , then the Lie algebra of  $T$ ,  $\mathfrak{t}$ , is a maximal abelian subalgebra of the Lie algebra of  $K$ ,  $\mathfrak{k}$ .

Using Eq. (2.4.103), it can be shown that the highest weight representation theorem for semisimple Lie algebras,  $\mathfrak{g}$ , (Theorem 2.4.49) extends to a highest weight representation theorem for compact connected Lie groups,  $K$  [Hal15, Thm. 12.6].

**Theorem 2.4.55.** *Let  $K$  be a connected, compact matrix Lie group and  $T$  a fixed maximal torus in  $K$ . Then*

1. every irreducible representation of  $K$  has a highest weight,
2. two irreducible representations of  $K$  with the same highest weight are isomorphic,
3. the highest weight of each irreducible representation is dominant and analytically integral, and
4. for every dominant, analytically integral element,  $\lambda$ , there exists an irreducible representation with highest weight  $\lambda$ .

*Proof.* See [Hal15, Sec. 12.5]. □

**Remark 2.4.56.** The discrepancy between the condition of being an *algebraically integral* weight (appearing in [Theorem 2.4.49](#)) and the stronger condition of being an *analytically integral* weight (appearing in [Theorem 2.4.55](#)) is a consequence of the fact that there can exist representations of semisimple Lie algebras,  $\mathfrak{g} = \mathfrak{k}_{\mathbb{C}}$ , which are not induced by a representation of compact connected Lie group with Lie algebra  $\mathfrak{k}$ . For instance ([\[Hal15, Ex. 12.11\]](#)), consider  $K = \mathrm{SO}(3)$  and  $\mathfrak{k} = \mathfrak{so}(3)$ , with basis

$$F_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \quad F_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \quad F_3 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad (2.4.112)$$

and maximal abelian subalgebra  $\mathfrak{t}$  spanned by the asymmetric matrix  $F_3$ . Then the only positive root is fixed by  $\alpha(F_3) = i$  (recall that  $\alpha$  is purely imaginary on  $\mathfrak{t}$  by [Lemma 2.4.34](#)) with root vector  $X = F_1 - iF_2 \in \mathfrak{so}(3)_{\mathbb{C}}$  because

$$[F_3, X] = [F_3, F_1] - i[F_3, F_2] = F_2 + iF_1 = iX = \alpha(F_3)X. \quad (2.4.113)$$

Since the associated coroot is

$$H_{\alpha} = \frac{2[X, X^*]}{\alpha([X, X^*])} = 2iF_3 \in i\mathfrak{t}, \quad (2.4.114)$$

a weight  $\lambda : i\mathfrak{t} \rightarrow \mathbb{R}$  is algebraically integral ( $\lambda(H_{\alpha}) = 2\lambda(iF_3) \in \mathbb{Z}$ ) if and only if  $\lambda(iF_3) \in \mathbb{R}$  is an integer or half integer:

$$2\lambda(iF_3) \in \mathbb{Z}. \quad (2.4.115)$$

On the other hand, as  $\exp(2\pi i\theta F_3) = \mathbb{1}_3$  is the identity in  $\mathrm{SO}(3)$  if and only if  $\theta \in \mathbb{Z}$ , a weight  $\lambda \in i\mathfrak{t} \rightarrow \mathbb{R}$  is analytically integral if and only if  $\lambda(iF_3)$  is an integer:

$$\lambda(iF_3) \in \mathbb{Z}. \quad (2.4.116)$$

To conclude this section, we introduce a definition of a *complex reductive group* which will be used throughout the thesis. Although the exact definition of a reductive group can be quite varied depending on the mathematical setting [[Lee01](#); [Mil14](#); [Kna01](#)], our definition is well-suited for our purposes. Just as we considered the complexification of a Lie algebra in [Section 2.4.5](#), there is a related notion of complexification for a compact Lie groups which relies on a universal construction. The *complexification* of a compact Lie group  $K$ , denoted  $G = K_{\mathbb{C}}$ , is a complex Lie group equipped with an inclusion map  $K \hookrightarrow G$  such that every smooth Lie group homomorphism from  $K$  to a complex Lie group  $Q$  lifts uniquely to a holomorphic homomorphism  $G \rightarrow Q$ . The key idea here is simply that representations of  $K$ ,  $\Phi : K \rightarrow \mathrm{GL}(\mathcal{V})$ , on complex vector spaces  $\mathcal{V}$ , gives rise to a representation  $\Phi : G \rightarrow \mathrm{GL}(\mathcal{V})$  of the larger group  $G$ .

**Definition 2.4.57.** A *complex reductive group* is a group  $G$  that is the complexification,  $G = K_{\mathbb{C}}$ , of a compact connected Lie group  $K$ . If  $\mathfrak{k}$  is the Lie algebra of  $K$ , then the Lie algebra of  $G$  is the complexification  $\mathfrak{g} = \mathfrak{k}_{\mathbb{C}} = \mathfrak{k} \oplus i\mathfrak{k}$ .



Since  $K$  is a compact and connected Lie group we have  $K = \exp(\mathfrak{k})$ . Similarly we can let  $\mathfrak{p} = i\mathfrak{k}$  and define a subset  $P := \exp(i\mathfrak{k})$  of  $G$ . Then the notion of a local Cartan involution,  $\theta : \mathfrak{g} \rightarrow \mathfrak{g}$ , on a complex reductive Lie algebra,  $\mathfrak{g}$  can be generalized to an involutive group isomorphism,  $\Theta : G \rightarrow G$ , on complex reductive group  $G$  called the **global Cartan involution**. This isomorphism  $\Theta$  fixes  $K \subset G$  and take inverses on  $P \subset G$ . As a quick summary,

$$\begin{aligned} A \in \mathfrak{k} &\implies \theta(A) = +A, & k \in K = \exp(\mathfrak{k}) &\implies \Theta(k) = k, \\ H \in \mathfrak{p} = i\mathfrak{k} &\implies \theta(H) = -H, & p \in P = \exp(i\mathfrak{k}) &\implies \Theta(p) = p^{-1}. \end{aligned} \quad (2.4.117)$$

In general, for any element  $g \in G$  of a complex reductive group, we define  $g^* \in G$  by

$$g^* := \Theta(g)^{-1}. \quad (2.4.118)$$

This notation is consistent with the previously defined \*-involution operation on  $X \in \mathfrak{g}$  by  $X^* = -\theta(X)$  in the sense that  $\exp(X^*) = \exp(X)^*$ . Furthermore, the universal property of the complexification of  $K$  means that any finite-dimensional unitary representation  $\Phi : K \rightarrow U(\mathcal{H})$  of  $K$  extends uniquely to a representation (also denoted by  $\Phi$ )  $\Phi : G \rightarrow GL(\mathcal{H})$  such that  $\Phi(g^*) = \Phi(g)^*$  where the latter \*-operation is simply the complex conjugate relative to the inner product on  $\mathcal{H}$ .

**Example 2.4.58.** Let  $d \in \mathbb{N}$  be a fixed finite dimension. For example if  $K = \mathrm{SU}(d)$  is the group of unitary matrices with determinant one, then the complexification of  $K = \mathrm{SU}(d)$  is  $G = \mathrm{SL}(d, \mathbb{C})$ , the group of complex matrices with determinant one. The Lie algebra of  $K = \mathrm{SU}(d)$  is  $\mathfrak{k} = \mathfrak{su}(d)$  the algebra of skew-Hermitian traceless matrices, while  $i\mathfrak{k} \simeq i\mathfrak{su}(d)$  is the algebra of Hermitian traceless matrices and  $P = \exp(i\mathfrak{su}(d))$  the set of positive definite matrices and  $\mathfrak{g} = \mathfrak{sl}(d, \mathbb{C})$  the Lie algebra of traceless matrices. In this example, the \*-involution on either  $X \in \mathfrak{sl}(d, \mathbb{C})$  or  $g \in \mathrm{SL}(d, \mathbb{C})$  is complex conjugation of matrices.

**Remark 2.4.59.** In general, the group multiplication in  $G$  gives rise to a diffeomorphism (but not group isomorphism)  $G \simeq K \times P$  called the **Cartan decomposition**. Specifically, as  $P := \exp(\mathfrak{p}) = \exp(i\mathfrak{k})$ , there exists an invertible map  $K \times \mathfrak{p} \rightarrow G$  given by

$$(k, X) \mapsto k \cdot \exp(X). \quad (2.4.119)$$

In particular, we have the following relationships between  $k, X$  and  $g$ :

$$\begin{aligned} g &= k \cdot \exp(X), \\ g^{-1} &= \exp(-X) \cdot k^{-1}, \\ \Theta(g) &= k \cdot \exp(-X), \\ g^* &= \exp(X) \cdot k^{-1}, \end{aligned} \quad (2.4.120)$$

Therefore,  $X$  and  $k$  are uniquely determined by  $g$  since

$$g^* \cdot g = \exp(2X) \implies \exp(X) = \sqrt{g^*g}, \quad (2.4.121)$$

and thus

$$k = \Theta(g)\sqrt{g^*g}. \quad (2.4.122)$$



When  $K = U(1)$  and thus  $G = \mathbb{C}_\times$ , this diffeomorphism states that every non-zero complex number  $z \in \mathbb{C}_\times$  can be written in polar form,  $z = e^{i\theta} |z|$  where  $e^{i\theta} \in U(1)$  and  $|z| > 1$  is a positive real number. More generally, if  $K = U(d)$  and thus  $G = GL(d, \mathbb{C})$ , this diffeomorphism states that every  $d \times d$  invertible complex matrix  $M \in GL(d, \mathbb{C})$  can be written uniquely in the form  $M = UP$  where  $U$  is a unitary matrix and  $P$  is a positive definite matrix.

A more detailed decomposition of elements in a complex reductive group is offered by the *Iwasawa decomposition* of  $G$ . Consider a generic complex reductive group,  $G = K_{\mathbb{C}}$ , which is the complexification of a connected compact Lie group  $K$ . A **Borel subgroup**  $B \subseteq G$  is a maximal solvable subgroup (which are all conjugate to each other by some elements of  $K$ ). Once a particular Borel subgroup  $B \subseteq G$  is chosen, its intersection with  $K$ ,

$$T = B \cap K, \quad (2.4.123)$$

constitutes a maximal torus in  $K$ . Moreover, the commutator subgroup of  $B$ ,

$$N = [B, B], \quad (2.4.124)$$

is a maximal unipotent subgroup of  $B$  and moreover  $T$  normalizes  $N$  inside  $B$ . The Lie algebras of  $B, N$  and  $T$  are denoted by  $\mathfrak{b}, \mathfrak{n}$  and  $\mathfrak{t}$ . Moreover, let  $\mathfrak{a} := i\mathfrak{t}$  and define  $A$  as the image under the exponential map of  $\mathfrak{a}$  in  $G$ .

**Proposition 2.4.60.** *Let  $G = K_{\mathbb{C}}$  be a complex reductive group with fixed Borel subgroup  $B \subseteq K$ . Then every group element  $g \in G$  can be uniquely expressed as*

$$g = k_g \cdot a_g \cdot n_g, \quad (2.4.125)$$

for unique  $k_g \in K$ ,  $a_g \in A$  and  $n_g \in N$ . This decomposition of  $g \in G$  is known as the **Iwasawa decomposition** [Iwa49].

**Example 2.4.61.** For the special case of  $K = U(d)$  and  $G = GL(d, \mathbb{C})$ , one can choose the Borel subgroup  $B$  to be the subgroup of upper triangular matrices and then  $T$  the subgroup of diagonal matrices. The maximal unipotent subgroup  $N$  of  $B$  is then the upper triangular matrices with ones along the main diagonal and the Iwasawa decomposition is essentially the QR decomposition of an invertible matrix where the diagonal part of the upper triangular matrix is factored out. When  $d = 3$  the Iwasawa decomposition can be expressed as

$$\underbrace{\begin{pmatrix} g_{11} & g_{12} & g_{13} \\ g_{21} & g_{22} & g_{23} \\ g_{31} & g_{32} & g_{33} \end{pmatrix}}_g = \underbrace{\begin{pmatrix} k_{11} & k_{12} & k_{13} \\ k_{21} & k_{22} & k_{23} \\ k_{31} & k_{32} & k_{33} \end{pmatrix}}_k \underbrace{\begin{pmatrix} a_1 & 0 & 0 \\ 0 & a_2 & 0 \\ 0 & 0 & a_3 \end{pmatrix}}_a \underbrace{\begin{pmatrix} 1 & n_{12} & n_{13} \\ 0 & 1 & n_{23} \\ 0 & 0 & 1 \end{pmatrix}}_n, \quad (2.4.126)$$

where  $g \in GL(3, \mathbb{C})$  is any invertible  $3 \times 3$  complex matrix,  $k \in U(3)$  is a unitary  $3 \times 3$  matrix, the diagonal of  $a$  is strictly positive  $(a_1, a_2, a_3) \in \mathbb{R}_{>0}^3$ , and  $(n_{12}, n_{13}, n_{23}) \in \mathbb{C}^3$ .

## 2.5 Glossary of Notation

Throughout this thesis, all logarithms are taken to be base  $e$ .

Table 2.1: Notation relating to Hilbert spaces.

$\mathcal{H}$	a complex inner product space
$\langle \cdot, \cdot \rangle$	the (sesquilinear) inner product associated to $\mathcal{H}$
$\ \cdot\ $	the norm induced by the $\langle \cdot, \cdot \rangle$
$\mathcal{H}_\times$	the non-zero elements of $\mathcal{H}$
$\mathcal{H}^*$	the dual vector space of linear functions on $\mathcal{H}$
$\text{End}(\mathcal{H})$	all (linear) operators acting on $\mathcal{H}$
$\mathcal{B}(\mathcal{H})$	bounded operators on $\mathcal{H}$
$\mathcal{B}_{\geq 0}(\mathcal{H})$	positive semi-definite operators on $\mathcal{H}$
$\text{GL}(\mathcal{H})$	invertible linear operators on $\mathcal{H}$
$\text{U}(\mathcal{H})$	unitary operators on $\mathcal{H}$
$\mathbb{P}\mathcal{H}$	the projective space associated to $\mathcal{H}$
$\psi$	a generic ray in the projective space $\mathbb{P}\mathcal{H}$
$P_\psi$	the rank-one projection operator onto the ray $\psi$
$\rho$	a generic density operator
$\varphi$	the functional form of a state on $\mathcal{H}$ ( $\varphi : \mathcal{B}(\mathcal{H}) \rightarrow \mathbb{C}$ )
$\mathcal{S}(\mathcal{H})$	the set of all density operators on $\mathcal{H}$

Table 2.2: Notation relating to probability measures.

$X$	a topological space (typically a Polish space)
$\Sigma(X)$	the Borel $\sigma$ -algebra on $X$
$\Delta$	a generic element in the Borel sigma algebra $\Sigma(X)$
$(X, \Sigma(X))$	a standard Borel space for $X$
$\mu$	a probability measure $\mu : \Sigma(X) \rightarrow [0, 1]$ on $X$
$\mathcal{P}(X)$	the set of all probability measures on $X$
$f_*\mu$	the push-forward of a measure $\mu$ through a measurable function $f$
$I$	a rate function for $X$ ( $I : X \rightarrow [0, \infty]$ )

Table 2.3: Notation relating to Lie groups, Lie algebras and representations.

$K$	a compact, sometimes connected, Lie group
$k$	a generic group element $k \in K$
$\mathfrak{k}$	the Lie algebra of the Lie group $K$
$G$	the complexification of $K$ ( $G = K_{\mathbb{C}}$ )
$g$	a generic group element $g \in G$
$\mathfrak{g}$	the Lie algebra of the complexification of $K$ ( $\mathfrak{g} = \mathfrak{k} \oplus i\mathfrak{k}$ )
$[\cdot, \cdot]$	the Lie bracket of a Lie algebra $\mathfrak{g}$
$\Theta$	the global Cartan involution on $G$ ( $\Theta : G \rightarrow G$ )
$\theta$	the local Cartan involution on $\mathfrak{g}$ ( $\theta : \mathfrak{g} \rightarrow \mathfrak{g}$ )
$X^*$	the star involution on $\mathfrak{g}$ ( $X^* = -\theta(X)$ )
$g^*$	the star involution on $G$ ( $g^* = \Theta(g)^{-1}$ )
$\Phi$	a group representation of $G$ on $\mathcal{H}$ ( $\Phi : G \rightarrow \text{GL}(\mathcal{H})$ )
$\phi$	a Lie algebra representation of $\mathfrak{g}$ on $\mathcal{H}$ ( $\phi : \mathfrak{k} \rightarrow \mathfrak{gl}(\mathcal{H})$ )
$\Pi_{\Phi}^{\lambda}$	subspace of highest weight vectors with weight $\lambda$ for $\Phi$
$\Pi_{\Phi}^{\lambda, k}$	the projection operator $\Pi_{\Phi}^{\lambda}$ conjugated by $k$
$\tilde{\Pi}_{\Phi}^{\lambda}$	the isotypic subspace of highest weight $\lambda$ for $\Phi$
$\text{cap}_{\Phi}$	the capacity map $\text{cap}_{\Phi} : \mathcal{H} \rightarrow [0, \infty]$ for $\Phi$
$\Omega_{\Phi}$	the moment map $\Omega_{\Phi} : \mathbb{P}\mathcal{H} \rightarrow i\mathfrak{k}^*$ for $\Phi$

Table 2.4: Commonly used symbols for the complexification of Lie algebras.

Group	Algebra	Relationship(s)	Comment
$K$	$\mathfrak{k}$		a compact Lie group
$G$	$\mathfrak{g}$	$\mathfrak{g} := \mathfrak{k} \oplus i\mathfrak{k}$	the complexification of $K$
$T$	$\mathfrak{t}$	$\mathfrak{t} \subseteq \mathfrak{k} \subseteq \mathfrak{g}$	the maximal torus of $K$
$H$	$\mathfrak{h}$	$\mathfrak{h} := \mathfrak{t} \oplus i\mathfrak{t}$	$\mathfrak{h}$ a Cartan subalgebra of $\mathfrak{g}$
$P$	$\mathfrak{p}$	$\mathfrak{p} := i\mathfrak{k}$	
$B$	$\mathfrak{b}$	$T = B \cap K$	a maximal solvable (Borel) subgroup of $G$
$N$	$\mathfrak{n}$	$N = [B, B] \subseteq B$	the maximal unipotent subgroup
$A$	$\mathfrak{a}$	$\mathfrak{a} := i\mathfrak{t}$	

# Chapter 3

## Non-commutative optimization

### 3.1 Introduction

In [Chapter 4](#), we will be interested in subject of quantum tomography which aims to estimate the properties of an unknown quantum system empirically by performing measurements on a large number of identical copies that system. Consequently, it will be useful to develop mathematical techniques to study probabilities of the form,

$$p_n(x | \rho) = \text{Tr}(\rho^{\otimes n} E_n^x), \quad (3.1.1)$$

where  $n$  is a large positive integer,  $\rho$  is a density operator acting on a Hilbert space  $\mathcal{H}$  describing an unknown quantum state, and  $E_n^x$  a positive semidefinite operator acting on  $\mathcal{H}^{\otimes n}$  describing the event of obtaining the estimate  $x$  upon performing a collective measurement on  $n$  identical copies of  $\rho$ . The purpose of this chapter, therefore, is to develop these mathematical methods using tools from invariant theory to study the large  $n$  limit of probabilities of the above form.

Before doing so, here we aim to provide a sketch for how the representation theory of non-compact groups will be used in this context. Consider a non-compact group  $G$  together with a non-unitary representation,  $\Phi : G \rightarrow \text{GL}(\mathcal{H})$ , of  $G$  on the Hilbert space  $\mathcal{H}$ . Now let a group element,  $g \in G$ , act on the density operator,  $\rho$ , by conjugation, sending  $\rho$  to the operator  $g \cdot \rho$ , defined by

$$g \cdot \rho := \Phi(g)\rho\Phi(g)^*, \quad (3.1.2)$$

where  $\Phi(g)^*$  is the complex conjugate of the operator  $\Phi(g)$ . Similarly, let  $g \in G$  act on the measurement effect  $E_n^x$  via

$$g \cdot E_n^x := \Phi^{\otimes n}(g)^* E_n^x \Phi^{\otimes n}(g). \quad (3.1.3)$$

If, in addition, there happens to exist a function,  $\chi_x : G \rightarrow (0, \infty)$ , independent of  $n$ , such that

$$g \cdot E_n^x \leq (\chi_x(g^{-1}))^n \mathbb{1}^{\otimes n}, \quad (3.1.4)$$

then a quick calculation reveals an upper bound on  $p_n(x | \rho)$  which holds for all  $g \in G$ :

$$p_n(x | \rho) = \text{Tr}(\rho^{\otimes n} E_n^x) = \text{Tr}((g \cdot \rho)^{\otimes n} (g^{-1} \cdot E_n^x)) \leq [\chi_x(g) \text{Tr}(g \cdot \rho)]^n. \quad (3.1.5)$$

Therefore, the likelihood of producing the estimate  $x$  decays at an exponential rate with respect to increasing  $n$ ,

$$p_n(x \mid \rho) \leq \exp(-nI_\rho(x)), \quad (3.1.6)$$

where the rate  $I_\rho(x) \in [0, \infty]$  is obtained by optimizing over all  $g \in G$ :

$$I_\rho(x) = -\log \inf_{g \in G} \chi_x(g) \text{Tr}(g \cdot \rho). \quad (3.1.7)$$

From the perspective of property tomography, therefore, it becomes desirable to find a sequence of measurements such that a quantum state,  $\rho$ , has vanishing rate,  $I_\rho(x) = 0$ , if and only if, the state  $\rho$  has property  $x$ . Finding examples of such measurements schemes will be the subject of [Chapter 4](#).

In this chapter, we will investigate a general class of optimization problems over a non-compact, non-commutative group  $G$ , similar to the one defined above. Our focus here will be on the general theory of non-commutative optimization from the perspective of geometric invariant theory; for instance the Kempf-Ness theorem provides a correspondence between extremal values and the vanishing of a generalized gradient, called the moment map [Theorem 3.2.31](#). Algorithmic implementations and complexity theoretic aspects of non-commutative optimization theory can be found elsewhere [\[Bür+19\]](#). The final result of this chapter, covered in [Section 3.3.4](#), is the *strong duality* result of Ref. [\[FW20\]](#), which serves as the foundation for many of the results in subsequent chapters.

## 3.2 Invariants & norm minimization

### 3.2.1 Group orbits & stability

The purpose of this section is to introduce the concept of group orbits associated to group representations and to study their topological closures. In particular, various notions of stability, borrowed from the subject of geometric invariant theory are directly related to whether or not the orbit of a vector under the action of a group is closed.

**Definition 3.2.1.** Let  $\mathcal{H}$  be a complex finite-dimensional Hilbert space and let  $\Phi : G \rightarrow \text{GL}(\mathcal{H})$  be a representation of a reductive group  $G$  such that the inner product  $\langle \cdot, \cdot \rangle$  on  $\mathcal{H}$  is invariant under the action of the maximal compact subgroup  $K \subseteq G$ . Let  $v \in \mathcal{H}$  be a vector in  $\mathcal{H}$ . The *orbit of  $v$* ,  $G \cdot v$ , is the set

$$G \cdot v = \Phi(G)v := \{\Phi(g)v \in \mathcal{H} \mid g \in G\}. \quad (3.2.1)$$

The *closure* of the orbit  $G \cdot v$ , denoted by  $\overline{G \cdot v}$  or  $\overline{\Phi(G)v}$ , is the topological closure of  $G \cdot v$  with respect to the norm-induced topology on  $\mathcal{H}$ .

**Example 3.2.2.** As representations are linear group actions, the orbit of the origin  $0 \in \mathcal{H}$  is the singleton set  $\Phi(G)0 := \{0\}$ . This orbit will be called the *trivial orbit*, while all other orbits are considered *non-trivial orbits*.

A particularly useful tool for verifying that two points belong to distinct orbits are invariant polynomials. We begin by recalling that the **ring of complex polynomials**, denoted  $\mathbb{C}[\mathcal{V}]$ , over a finite-dimensional complex vector space  $\mathcal{V}$ , can be defined either in a coordinate-full or coordinate-free manner. The coordinate-free definition for  $\mathbb{C}[\mathcal{V}]$  proceeds by considering the commutative ring of functions from  $\mathcal{V}$  to  $\mathbb{C}$  generated by elements of the dual space  $\mathcal{V}^*$  of  $\mathbb{C}$ -linear functions on  $\mathcal{V}$ . From this perspective,  $\mathbb{C}[\mathcal{V}]$  is also sometimes called the symmetric algebra on  $\mathcal{V}$ . Alternatively, a coordinate-full definition proceeds by identifying some basis  $\{e_1, \dots, e_d\}$  for  $\mathcal{V}$  with the indeterminate variables  $\{x_1, \dots, x_d\}$  such that the polynomials in  $\mathbb{C}[\mathcal{V}]$  can be identified with the set of complex polynomials in  $d$  variables, denoted  $\mathbb{C}[x_1, \dots, x_d]$ . In either case, any representation of a group  $G$  on  $\mathcal{V}$  induces an action of  $G$  on  $\mathbb{C}[\mathcal{V}]$  and subsequently gives rise to the notion of an invariant polynomial.

**Definition 3.2.3.** Let  $\mathcal{V}$  be a finite-dimensional complex vector space and  $\mathbb{C}[\mathcal{V}]$  the complex polynomial ring on  $\mathcal{V}$ . Given a representation  $\Phi : G \rightarrow \text{GL}(\mathcal{V})$  of a group  $G$  on  $\mathcal{V}$ , there is a natural action of the group  $G$  on  $\mathbb{C}[\mathcal{V}]$ , denoted here by  $\Phi_*$ , and defined for  $p \in \mathbb{C}[\mathcal{V}]$  and  $g \in G$  by

$$(\Phi_*(g)(p))(v) := p(\Phi(g^{-1})v). \quad (3.2.2)$$

A polynomial  $p \in \mathbb{C}[\mathcal{V}]$  is said to be a **invariant** (or sometimes  $G$ -invariant) polynomial if for all  $g \in G$ ,  $\Phi_*(g)(p) = p$ . The subset of all invariant polynomials forms a subring of  $\mathbb{C}[\mathcal{V}]$ , called the **ring of invariant polynomials** and is denoted by  $\mathbb{C}[\mathcal{V}]^\Phi$ .

**Remark 3.2.4.** Recall that in this setting, every group  $G$  acts *linearly* on the vector space  $\mathcal{V}$  through the representation  $\Phi : G \rightarrow \text{GL}(\mathcal{H})$ . Consequently, the degree of a homogeneous polynomial  $p \in \mathbb{C}[\mathcal{V}]$  is always preserved by the action of  $G$  meaning,

$$\forall g \in G : \deg(\Phi_*(g)p) = \deg(p). \quad (3.2.3)$$

Therefore, the ring of  $G$ -invariant polynomials,  $\mathbb{C}[\mathcal{V}]^\Phi$ , naturally forms a *graded ring* in the sense that

$$\mathbb{C}[\mathcal{V}]^\Phi = \bigoplus_{n=0}^{\infty} \mathbb{C}[\mathcal{V}]_n^\Phi \quad (3.2.4)$$

where each summand,  $\mathbb{C}[\mathcal{V}]_n^\Phi$ , is the vector space of homogeneous degree- $n$   $G$ -invariant polynomials, and furthermore for every  $n, m \in \mathbb{N}$ , the product of a degree- $n$  and degree- $m$   $G$ -invariant polynomial is a degree- $(n + m)$   $G$ -invariant polynomial:

$$\mathbb{C}[\mathcal{V}]_n^\Phi \otimes \mathbb{C}[\mathcal{V}]_m^\Phi \subseteq \mathbb{C}[\mathcal{V}]_{n+m}^\Phi. \quad (3.2.5)$$

**Remark 3.2.5.** The definition of a  $G$ -invariant polynomial  $p \in \mathbb{C}[\mathcal{V}]^\Phi$  immediately implies that  $p$  takes constant values when evaluated on orbits because  $p(\Phi(g)v) = p(v)$  for all  $g \in G$ . Moreover, since polynomials are continuous functions, they are also constant when evaluated on the closures of orbits, i.e., for all  $w \in \overline{\Phi(g)v}$ , we have  $p(w) = p(v)$  also. The contrapositive version of this observation is quite useful; if  $v, w \in \mathcal{V}$  are vectors such that  $p(v) \neq p(w)$  for some  $G$ -invariant polynomial  $p$ , one concludes that  $v$  and  $w$  must belong to distinct orbit closures, i.e., the intersection of their orbit closures is empty:  $\overline{\Phi(g)v} \cap \overline{\Phi(g)w} = \emptyset$ .

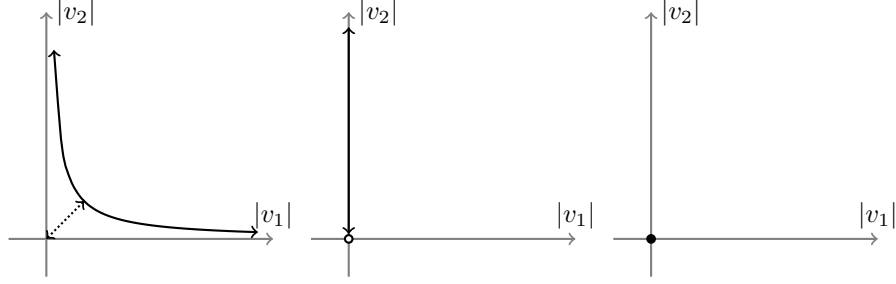


Figure 3.1: Three distinct types of group orbits under the group action sending the pair of complex numbers  $(v_1, v_2)$  to  $z \cdot (v_1, v_2) = (zv_1, z^{-1}v_2)$ , visualized using the moduli of the complex coordinates,  $(|v_1|, |v_2|)$ .

**Example 3.2.6.** Perhaps the simplest non-trivial example of an invariant polynomial arises under the action of the finite multiplicative group  $\mathbb{Z}_2 = \{-1, +1\}$  on the one-dimensional complex vector space  $\mathcal{V} \cong \mathbb{C}$  where  $-1 \in \mathbb{Z}_2$  acts on  $x \in \mathbb{C}$  by reflection  $x \mapsto -x$ . In this example, the polynomial ring  $\mathbb{C}[\mathcal{V}] \simeq \mathbb{C}[x]$  is univariate and consists of polynomials of the form  $c_0 + c_1x + c_2x^2 + \dots$ . Under the action of  $\mathbb{Z}_2$  by  $x \mapsto -x$ , the invariant polynomials  $\mathbb{C}[x]^{\mathbb{Z}_2}$  are precisely those containing only even-degree terms, i.e.  $c_0 + c_2x^2 + \dots$ .

**Example 3.2.7.** For another example involving a continuous group, consider the two-dimensional vector space  $\mathcal{V} = \mathbb{C}^2$  and the multiplicative group of non-zero complex numbers  $G = \mathbb{C}_\times$  acting on  $(v_1, v_2) \in \mathbb{C}^2$  by sending  $z \in \mathbb{C}_\times$  to the linear transformation  $\Phi(z) \in \text{End}(\mathbb{C}^2)$  defined by

$$(v_1, v_2) \mapsto (zv_1, z^{-1}v_2). \quad (3.2.6)$$

The orbits under the above action fall into one of three qualitatively distinct categories (depicted in Fig. 3.1) which can be determined by considering a representative element  $(v_1, v_2)$  of the orbit:

- i) If  $(v_1, v_2) = (0, 0)$ , then the orbit is simply the trivial orbit  $\Phi(\mathbb{C}_\times)(0, 0) = \{(0, 0)\}$ .
- ii) If  $v_1 = a \neq 0$  but  $v_2 = 0$ , then the orbit consists of the  $v_2 = 0$  axis in  $\mathbb{C}^2$  where the origin has been removed  $\Phi(\mathbb{C}_\times)(a, 0) = \{(z, 0) \mid z \in \mathbb{C}_\times\} \simeq \mathbb{C}_\times$  (or analogously if  $v_1 = 0$  and  $v_2 = b \neq 0$ ).
- iii) Or if  $(v_1, v_2) = (a, b)$  are both non-zero, then the orbit consists of all points such that  $v_1v_2 = ab$ , i.e.  $\Phi(\mathbb{C}_\times)(a, b) = \{(v_1, v_2) \mid v_1v_2 = ab\}$ .

Moreover, the ring of invariant polynomials,  $\mathbb{C}[v_1, v_2]^{\mathbb{C}_\times}$ , is generated from a single invariant polynomial of homogeneous degree two, namely  $v_1v_2$ . Indeed the third type of orbit above is completely characterized as the solution set to the polynomial equation  $v_1v_2 = c$  where  $c \in \mathbb{C}$  is a non-zero complex constant. By comparison, it can be shown that the orbits of type i) or ii) cannot be expressed as the solution set of invariant polynomials in  $\mathbb{C}[v_1, v_2]^{\mathbb{C}_\times}$ .

In the previous example (Example 3.2.7), it was shown that the orbits of a group acting on a vector space cannot always be characterized as the solutions to a family of invariant polynomials. In general, any non-zero group orbit whose topological closure contains the

origin of the vector space cannot be separated from the zero orbit using invariant polynomials invariant (or more generally invariant continuous functions). Therefore, distinguishing between different group orbits is highly sensitive to whether or not the topological closure of the group orbit contains the origin. This observation gives rise to the notion of stability that will be used subsequently.

**Definition 3.2.8.** Here  $\mathcal{V}$  will be a complex finite-dimensional Hilbert space, and  $\Phi : G \rightarrow \text{GL}(\mathcal{V})$  a representation of a reductive group  $G$  acting on  $\mathcal{V}$ . Let  $v \in \mathcal{V}$  be a fixed vector. Then  $v$  is said to be

- i) **unstable** if  $0 \in \overline{\Phi(G)v}$ ,
- ii) **semistable** if  $0 \notin \overline{\Phi(G)v}$ ,
- iii) **polystable** if  $v$  is semistable and  $\Phi(G)v$  is closed, i.e.  $\Phi(G)v = \overline{\Phi(G)v}$ ,
- iv) and **stable** if  $v$  is polystable and additionally the stabilizer of  $v$ ,  $G_v$ , has finite order.

**Remark 3.2.9.** Linguistically, one might expect the notion of a *stable* vector to be any vector that is *not* unstable. However, this is not the case; the logical opposite of an unstable vector is the notion of a *semistable* vector.

**Example 3.2.10.** Returning to [Example 3.2.7](#) and consulting [Fig. 3.1](#), one can see that with respect to the operation  $(v_1, v_2) \mapsto (zv_1, z^{-1}v_2)$  for  $z \in \mathbb{C}_\times$ , vectors  $(v_1, v_2)$  lying on either the  $v_1$  or  $v_2$  axes are unstable because their orbit closures contain the origin. By comparison, if  $(v_1, v_2)$  does not lie on either  $v_1$  or  $v_2$  axes, then  $(v_1, v_2)$  is in fact semistable because their orbit closure excludes the origin. Moreover, in this example, the semistable points are additionally polystable (because their orbits are topologically closed) and stable (because their stabilizer is the trivial group and has order one which is finite).

These examples bring us to a fundamental result of geometric invariant theory which can be seen as a strengthening of the link between the intersections of orbit closures and invariant polynomials previously mentioned in [Remark 3.2.5](#).

**Proposition 3.2.11.** *Let  $G$  be a complex reductive group, let  $\Phi : G \rightarrow \text{GL}(\mathcal{V})$  be a rational representation of  $G$  on a complex finite-dimensional vector space  $\mathcal{V}$ , and fix  $v, w \in \mathcal{V}$ . Then  $\overline{\Phi(G)v} \cap \overline{\Phi(G)w} \neq \emptyset$  if and only if  $p(v) = p(w)$  for all invariant polynomials  $p \in \mathbb{C}[\mathcal{V}]^\Phi$ .*

If one of the two vectors in [Proposition 3.2.11](#) is taken to be the zero vector, one obtains the following corollary which relates the semistability of a vector ([Definition 3.2.8](#)) to the values it takes on homogeneous invariant polynomials.

**Corollary 3.2.12.** *Let  $\Phi : G \rightarrow \text{GL}(\mathcal{V})$  be as in [Proposition 3.2.11](#) and fix  $v \in \mathcal{V}$ . Then  $0 \in \overline{\Phi(G)v}$  if and only if  $p(v) = p(0)$  for all invariant polynomials  $p \in \mathbb{C}[\mathcal{V}]$ . Equivalently,  $v$  is unstable if and only if  $p(v) = 0$  for all homogeneous invariant polynomials  $p \in \bigoplus_{n=1}^{\infty} \mathbb{C}[\mathcal{V}]_n^\Phi$ .*

**Definition 3.2.13.** The **null cone**, denoted by  $\mathcal{N}$ , is the set of all vectors  $v \in \mathcal{V}$  which cannot be separated from the origin by invariant polynomials, i.e.

$$\mathcal{N} := \{v \in \mathcal{H} \mid 0 \in \overline{\Phi(G)v}\}. \quad (3.2.7)$$



### 3.2.2 Invariant and fixed subspaces

Recall, from [Definition 2.4.5](#), that a subspace  $\mathcal{W} \subseteq \mathcal{V}$  is invariant under the action of a group representation  $\Phi : G \rightarrow \text{GL}(\mathcal{V})$  if for all group elements  $g \in G$ , the linear operation,  $\Phi(g)$ , maps every vector  $w \in \mathcal{W}$  inside  $\mathcal{W}$  to another vector  $\Phi(g)w \in \mathcal{W}$  inside  $\mathcal{W}$ . There is a stronger notion of invariance for subspaces that requires each *vector*  $w \in \mathcal{W}$  itself to be invariant. This stronger notion of invariance will be important in [Section 3.3.3](#). Note that in some references, e.g. [\[FW20\]](#), this subspace of invariant vectors is referred to as *the* invariant subspace. In this thesis, we instead prefer to use the terminology of the *fixed subspace*.

**Definition 3.2.14.** Let  $\Phi : G \rightarrow \text{GL}(\mathcal{H})$  be a representation of a group  $G$  on a complex finite-dimensional Hilbert space  $\mathcal{H}$ . The **fixed subspace** is the subspace  $\mathcal{H}^\Phi \subseteq \mathcal{H}$  consisting of all vectors on which  $\Phi$  acts trivially:

$$\mathcal{H}^\Phi := \{v \in \mathcal{H} \mid \forall g \in G, \Phi(g)v = v\}. \quad (3.2.8)$$

The fixed subspace  $\mathcal{H}^\Phi$  can also be identified with the multiplicity space for the trivial subrepresentation of  $\Phi$ , meaning

$$\mathcal{H}^\Phi \simeq \text{Hom}_G(1_G, \Phi) \quad (3.2.9)$$

where the symbol,  $1_G$ , here denotes the trivial representation of  $G$  sending  $g \in G$  to the multiplicative unit  $1_G(g) = 1 \in \mathbb{C}$ .

**Definition 3.2.15.** Let  $\Phi : G \rightarrow \text{GL}(\mathcal{H})$  be a representation of a group  $G$  on a complex finite-dimensional Hilbert space  $\mathcal{H}$ . The **fixed subspace of degree  $n$**  is fixed subspace  $(\mathcal{H}^{\otimes n})^\Phi \subseteq \mathcal{H}^{\otimes n}$  of the  $n$ th tensor power representation  $\Phi^{\otimes n}$ :

$$(\mathcal{H}^{\otimes n})^\Phi := \{w \in \mathcal{H}^{\otimes n} \mid \forall g \in G, \Phi^{\otimes n}(g)w = w\}. \quad (3.2.10)$$

The projection operator onto the fixed subspace of degree  $n$ ,  $(\mathcal{H}^{\otimes n})^\Phi \subseteq \mathcal{H}^{\otimes n}$ , will always be denoted by  $\Gamma_{\Phi^{\otimes n}}$ .

**Remark 3.2.16.** For any positive integers  $n, m \in \mathbb{N}$ , there exists a map sending  $w_n \in (\mathcal{H}^{\otimes n})^\Phi$  and  $w_m \in (\mathcal{H}^{\otimes m})^\Phi$  to the vector

$$w_n \otimes w_m \in (\mathcal{H}^{\otimes(n+m)})^\Phi, \quad (3.2.11)$$

and through this map,  $(\mathcal{H}^{\otimes n})^\Phi \otimes (\mathcal{H}^{\otimes m})^\Phi$  can be viewed as a subspace of  $(\mathcal{H}^{\otimes(n+m)})^\Phi$ :

$$(\mathcal{H}^{\otimes n})^\Phi \otimes (\mathcal{H}^{\otimes m})^\Phi \subseteq (\mathcal{H}^{\otimes(n+m)})^\Phi. \quad (3.2.12)$$

If  $\Gamma_{\Phi^{\otimes n}} \in \text{End}(\mathcal{H}^{\otimes n})$  is the projection operator onto the subspace  $(\mathcal{H}^{\otimes n})^\Phi \subset \mathcal{H}^{\otimes n}$ , then we obtain the operator inequality

$$\Gamma_{\Phi^{\otimes n}} \otimes \Gamma_{\Phi^{\otimes m}} \leq \Gamma_{\Phi^{\otimes(n+m)}}. \quad (3.2.13)$$

**Lemma 3.2.17.** *Let  $G$  be a complex reductive group which is the complexification of a compact, connected Lie group  $K$ , i.e.  $G = K_{\mathbb{C}}$ . Let  $\mathcal{H}$  be a complex finite-dimensional Hilbert space and  $\Phi : G \rightarrow \mathrm{GL}(\mathcal{H})$  a representation of  $G$  on  $\mathcal{H}$ . Let  $\Gamma_{\Phi} \in \mathrm{End}(\mathcal{H})$  denote the projection operator onto the subspace  $\mathcal{H}^{\Phi} \subseteq \mathcal{H}$  fixed by  $G$ . Then  $\Gamma_{\Phi}$  can be expressed as an integral over the compact group  $K$ ,*

$$\Gamma_{\Phi} = \int_{k \in K} \Phi(k) \, d\mu(k), \quad (3.2.14)$$

where  $\mu : \Sigma(K) \rightarrow [0, 1]$  denotes the normalized  $K$ -invariant Haar measure.

*Proof.* From the  $K$ -invariance of the Haar measure  $du$  and Schur's lemma, one can conclude that the right-hand side of Eq. (3.2.14) is equal to the projector operator onto the subspace fixed by the restriction of  $\Phi$  onto  $K$ , denoted by  $\Phi|_K$ . To prove Lemma 3.2.17, it suffices to show that the subspaces fixed by  $G$  (via  $\Phi$ ) and  $K$  (via  $\Phi|_K$ ) indeed coincide. This follows by noticing that when  $v$  is fixed by  $K$ , for all  $X$  in the Lie algebra  $\mathfrak{k}$  of  $K$ ,  $\phi(X)v = 0$ . Moreover, since  $X \in \mathfrak{k}$  is arbitrary and  $G$  is the complexification of  $K$ ,  $\phi(X+iY)v = \phi(X)v + i\phi(Y)v = 0$  for all  $X, Y \in \mathfrak{k}$  and thus  $\phi(Z)v = 0$  holds for all  $Z \in \mathfrak{g} = \mathfrak{k} \oplus i\mathfrak{k}$  and thus  $v$  is also  $\Phi$ -invariant.  $\square$

**Corollary 3.2.18.** *Let  $\Phi : G \rightarrow \mathrm{GL}(\mathcal{H})$  be as in Lemma 3.2.17. Then the projection operator onto the fixed subspace of degree  $n$ ,  $(\mathcal{H}^{\otimes n})^{\Phi}$  may be expressed as*

$$\Gamma_{\Phi^{\otimes n}} = \int_{k \in K} \Phi^{\otimes n}(k) \, d\mu(k). \quad (3.2.15)$$

*Proof.* The claim follows by applying Lemma 3.2.17 to the  $n$ th tensor power representation  $\Phi^{\otimes n} : G \rightarrow \mathrm{GL}(\mathcal{H}^{\otimes n})$ .  $\square$

**Remark 3.2.19.** Note that the notion of an invariant polynomial of degree  $n$  (encountered in Definition 3.2.3) is intimately related to the notion of an element of the fixed subspace of degree  $n$ . Recall that a homogeneous degree  $n$  polynomial  $p \in \mathbb{C}[\mathcal{H}]_n$  is characterized by the fact that  $p(xv) = x^n p(v)$  for all  $x \in \mathbb{C}$ . Furthermore every homogeneous degree  $n$  polynomial,  $p \in \mathbb{C}[\mathcal{H}]_n$ , can be identified by its coefficients with respect to a basis of monomials (of degree  $n$ ), and thus it can be uniquely associated with an element of the  $n$ th symmetric subspace  $\mathrm{Sym}^n(\mathcal{H})$  in the sense that for each  $p \in \mathbb{C}[\mathcal{H}]_n$  there exists a  $w \in \mathrm{Sym}^n(\mathcal{H})$  such that  $p(v) = \langle w, v^{\otimes n} \rangle$  holds for all  $v \in \mathcal{H}$ .

**Remark 3.2.20.** For any representation  $\Phi : G \rightarrow \mathrm{GL}(\mathcal{H})$ , the representation  $\phi^{\otimes n} : G \rightarrow \mathrm{GL}(\mathcal{H})$  naturally commutes with the permutation of the tensor factors in  $\mathcal{H}^{\otimes n}$ . In view of Remark 3.2.19, the space of homogeneous degree  $n$ ,  $G$ -invariant polynomials,  $\mathbb{C}[\mathcal{H}]_n^{\Phi}$ , can therefore be identified with the space of symmetric,  $G$ -invariant vectors in  $\mathcal{H}^{\otimes n}$ . In particular, we have the following relationship:

$$\mathbb{C}[\mathcal{H}]_n^{\Phi} \simeq (\mathcal{H}^{\otimes n})^{\Phi} \cap \mathrm{Sym}^n(\mathcal{H}). \quad (3.2.16)$$

### 3.2.3 Capacities & moment maps

While [Section 3.2.1](#) was primarily concerned with whether the orbit of a vector  $v \in \mathcal{H}$  under the action of a group representation  $\Phi : G \rightarrow \text{GL}(\mathcal{H})$  contains the origin (or gets arbitrarily close to the origin), the purpose of this section is to introduce concepts which characterize these notions of stability more quantitatively.

Specifically, since the Hilbert space  $\mathcal{H}$  is equipped with a positive-definite norm  $\|\cdot\| : \mathcal{H} \rightarrow \mathbb{R}_{\geq 0}$ , the magnitude of  $\|\Phi(g)v\|$  as  $g$  varies through the group  $G$  serves as a measure of the distance between  $\Phi(g)v$  and the origin  $0 \in \mathcal{H}$ . By minimizing over the whole group  $G$  we obtain a measure of distance between the orbit  $\Phi(G)v$  and the origin called the *capacity of  $v$* .

**Definition 3.2.21.** Let  $\Phi : G \rightarrow \text{GL}(\mathcal{V})$  and let  $v \in \mathcal{V}$ . The *capacity of the vector  $v$*  is defined as

$$\text{cap}(v) := \inf_{g \in G} \|\Phi(g)v\| = \min_{w \in \overline{\Phi(G)v}} \|w\|. \quad (3.2.17)$$

**Example 3.2.22.** Again returning to [Example 3.2.7](#) and consulting [Fig. 3.1](#), one can see that the capacity of a vector  $(v_1, v_2) \in \mathbb{C}^2$  under the group acting  $(v_1, v_2) \mapsto (zv_1, z^{-1}v_2)$  for  $z \in \mathbb{C}_x$ , defined as

$$\text{cap}((v_1, v_2)) = \inf_{z \in \mathbb{C}_x} \|(zv_1, z^{-1}v_2)\| = \inf_{z \in \mathbb{C}_x} \sqrt{|zv_1|^2 + |z^{-1}v_2|^2}, \quad (3.2.18)$$

depends on which of the three types of orbits the vector  $(v_1, v_2)$  belongs. First, the capacity of the zero vector is evidently zero,  $\text{cap}((0, 0)) = 0$ , as the norm of the zero vector is always zero. Second, the capacity of any vector  $(v_1, v_2)$  lying on, say, the  $v_1 = 0$  axis has zero capacity because in the limit as  $|z|$  tends to infinity, the value of  $|z^{-1}v_2|$  tends to zero and thus for all  $v_2 \in \mathbb{C}$ ,  $\text{cap}((0, v_2)) = 0$ . Finally, the capacity of a vector  $(v_1, v_2)$  which satisfies both  $v_1 \neq 0$  and  $v_2 \neq 0$  has a positive capacity equal to

$$\text{cap}((v_1, v_2)) = \inf_{x \in \mathbb{R}_{>0}} \sqrt{x|v_1|^2 + x^{-1}|v_2|^2} = \sqrt{2|v_1||v_2|}, \quad (3.2.19)$$

and this value is attained when  $x = |z| = |v_2|/|v_1|$  (see [Fig. 3.1](#)).

**Remark 3.2.23.** In general, the value of the capacity of a vector  $v$  is evidently related to the stability properties of  $v$ . Indeed one can readily verify that  $\text{cap}(v) = 0$  if and only if  $v$  is unstable, or equivalently,  $\text{cap}(v) > 0$  if and only if  $v$  is semistable. The specific magnitude of the capacity when  $\text{cap}(v) > 0$  is therefore unimportant for the purposes of assessing stability. While the definition given above is consistent with references [\[Bür+19; FW20\]](#), some authors, e.g. [\[Amé+21a\]](#), prefer to define capacity as  $\inf_{g \in G} \|\Phi(g)v\|^2$ . The specific value of the capacity, beyond whether or not it vanishes, will become important in [Section 4.4.1](#).

**Remark 3.2.24.** Notice that by linearity of the representation  $\Phi : G \rightarrow \text{GL}(\mathcal{V})$ , the capacity of any vector  $v \in \mathcal{V}$  satisfies the following scaling property:

$$\forall z \in \mathbb{C} : |z| \text{cap}(v) = \text{cap}(zv). \quad (3.2.20)$$

Since this property holds for all groups  $G$ , it becomes reasonable to define the capacity for the ray  $\psi \in \mathbb{P}\mathcal{H}$ , termed the **projective capacity** of  $\psi$ , by

$$\text{cap}(\psi) := \inf_{g \in G} \text{Tr}(P_\psi \Phi(g^*g))^{\frac{1}{2}}. \quad (3.2.21)$$

The capacity of a non-zero vector,  $v \in \mathcal{V}_x$ , and the capacity of the ray  $[v] \in \mathbb{P}\mathcal{H}$  containing  $v$  are thus related by

$$\begin{aligned} \text{cap}([v]) &= \inf_{g \in G} \text{Tr}(P_{[v]} \Phi(g^*g))^{\frac{1}{2}}, \\ &= \|v\|^{-1} \inf_{g \in G} \sqrt{\langle v, \Phi(g^*g)v \rangle}, \\ &= \|v\|^{-1} \inf_{g \in G} \sqrt{\langle \Phi(g)v, \Phi(g)v \rangle}, \\ &= \|v\|^{-1} \inf_{g \in G} \|\Phi(g)v\|, \\ &= \|v\|^{-1} \text{cap}(v). \end{aligned} \quad (3.2.22)$$

Although the notion of projective capacity is well-defined and useful, the capacity of a vector,  $\text{cap}(v)$ , has the conceptually useful property that the infimum of  $\|\Phi(g)v\|$  over  $g \in G$  always holds for some vector  $w \in \overline{\Phi(G)v}$  in the orbit closure of  $v$  so that  $\text{cap}(v) = \|w\|$ , whereas the closure of the corresponding projective orbit is not well defined when  $\text{cap}(v) = 0$ .

**Remark 3.2.25.** For any vector  $v \in \mathcal{H}$ , the capacity  $\text{cap}(v)$  is lower-bounded by zero (as the norm is non-negative) and upper bounded by the norm  $\|v\|$  (because the identity element  $e \in G$  yields  $\|\Phi(e)v\| = \|v\|$ ):

$$0 \leq \text{cap}(v) \leq \|v\|. \quad (3.2.23)$$

If  $v$  happens to be such that  $\text{cap}(v) = \|v\|$ , then by definition,  $v$  attains the minimal norm in its orbit,  $\Phi(G)v$ , and thus the norm of  $\Phi(g)v$  is always equal to or greater than the norm of  $v$  itself, i.e.  $\|\Phi(g)v\| \geq \|v\|$  for all  $g \in G$ .

**Remark 3.2.26.** Consider the case where the group  $G$  is the complexification,  $G = K_{\mathbb{C}}$ , of a connected compact Lie group  $K$  (see Section 2.4), and additionally where the norm  $\|\cdot\| : \mathcal{H} \rightarrow \mathbb{R}_{\geq 0}$  on the representation space  $\mathcal{H}$  is  $K$ -invariant in the sense that  $\|\Phi(k)v\| = \|v\|$  for all  $k \in K$ . The  $K$ -invariance of the norm means that the function sending a group element  $g \in G$  to the norm  $\|\Phi(g)v\|$  is constant over left cosets  $Kg \in K \backslash G$  where

$$Kg := \{kg \in G \mid k \in K\}, \quad (3.2.24)$$

simply because for all elements  $h \in Kg$  of the left coset  $Kg$  satisfies

$$\|\Phi(h)v\| = \|\Phi(k)\Phi(g)v\| = \|\Phi(g)v\|. \quad (3.2.25)$$

Therefore, in computing the capacity of a vector  $v$ , one need not optimize  $\|\Phi(g)v\|$  as  $g$  varies over the whole group, but instead, one only needs to optimize  $\|\Phi(g)v\|$  as  $g$  varies over distinct representatives from the set of all left cosets, denoted by  $K \backslash G$ . Furthermore, recall from Section 2.4.7 that the Cartan decomposition  $G \simeq K \times P$  of  $G$  means the set of left cosets  $K \backslash G$  is diffeomorphic to the subset  $P \subseteq G$  in the sense that the left coset  $Kg$  is identified with the value of  $g^*g \in P$ . Therefore, capacity of a vector  $v$  can also be understood as the optimum of the function sending  $p \in P$  to  $\sqrt{\langle v, \Phi(p)v \rangle}$ .

**Remark 3.2.27.** The problem of computing the capacity of a vector  $v \in \mathcal{H}$  can be conceptualized as an optimization problem over the group  $G$  where the objective function whose output is being optimized is the function  $g \mapsto \|\Phi(g)v\|$  or equivalently, the function  $g \mapsto \log \|\Phi(g)v\|$ , also known as the *Kempf-Ness function*.

**Definition 3.2.28.** Let  $\Phi : G \rightarrow \mathrm{GL}(\mathcal{H})$  be a representation of a complex reductive group  $G$  on a complex finite-dimensional Hilbert space  $\mathcal{H}$ . Let  $v \in \mathcal{H}_\times$  be non-zero. The **Kempf-Ness function**  $F_v : G \rightarrow \mathbb{R}$  is defined for  $g \in G$  by

$$F_v(g) = \log \|\Phi(g)v\|. \quad (3.2.26)$$

**Remark 3.2.29.** In order to characterize the extremal points of the Kempf-Ness function  $F_v : G \rightarrow \mathbb{R}$ , it will be useful to consider the derivative of the function  $f_v : \mathbb{R} \rightarrow \mathbb{R}$  for  $Z \in \mathfrak{g}$  defined for  $t \in \mathbb{R}$  as follows, where  $Z \in \mathfrak{g}$  is an element of the Lie algebra of  $G$ , and

$$f(t) := F_v(\exp(tZ)) = \log \|\Phi(\exp(tZ))v\| = \log \|\exp(t\phi(Z))v\|. \quad (3.2.27)$$

When  $G$  is the complexification of a compact Lie group  $K$ , the Lie algebra of  $G$  is  $\mathfrak{g} = \mathfrak{k} \oplus i\mathfrak{k}$  where  $\mathfrak{k}$  is the Lie algebra of  $K$ . If, in addition, the norm  $\|\cdot\|$  is  $K$ -invariant, then  $f_v$  is a constant function in the direction  $X \in \mathfrak{k}$  corresponding to the Lie algebra of compact Lie group  $K$  (this statement is the equivalent of Remark 3.2.26 from the Lie algebraic perspective). Therefore, the Kempf-Ness function only changes along curves  $t \mapsto \exp(tX)$  where  $X$  belongs to the subspace  $i\mathfrak{k} \subset \mathfrak{g}$  (i.e., when  $X$  satisfies  $X^* = X$ ). In this setting, one obtains

$$f(t) := F_v(\exp(tX)) = \log \|\exp(t\phi(X))v\| = \log \|v_t\|. \quad (3.2.28)$$

where  $v_t = \exp(t\phi(X))v$ . Letting  $u_t$  denote the unit vector  $v_t/\|v_t\|$ , the first two derivatives of  $f(t)$  are [Bür+19, Eq. (3.13)]

$$f'(t) = \langle u_t, \phi(X)u_t \rangle, \quad \text{and} \quad (3.2.29)$$

$$f''(t) = 2(\langle \phi(X)u_t, \phi(X)u_t \rangle - \langle u_t, \phi(X)u_t \rangle^2) = \|u_t \wedge \phi(X)u_t\|^2, \quad (3.2.30)$$

where  $u_t \wedge \phi(X)u_t$  is the antisymmetric vector defined by

$$u_t \otimes \phi(X)u_t - \phi(X)u_t \otimes u_t. \quad (3.2.31)$$

There are two important conclusions to be drawn from the above calculation. First, the non-negativity of the second derivative of  $f$  demonstrates that  $f$  is a *convex* function. In the language of [Bür+19], this proves that the Kempf-Ness function,  $F_v$ , is geodesically convex. Second, the first derivative of  $f$  evaluated at  $t = 0$ ,  $f'(0)$ , when viewed as being functionally dependent on the value of  $X \in i\mathfrak{k}$ , gives rise to the notion of the moment map, defined below.

**Definition 3.2.30.** Let  $\Phi : G \rightarrow \mathrm{GL}(\mathcal{H})$  be the representation of reductive group  $G$  with maximal compact subgroup  $K$  and Lie algebra  $\mathfrak{g} = \mathfrak{k} \oplus i\mathfrak{k}$  and  $\phi : \mathfrak{g} \rightarrow \mathfrak{gl}(\mathcal{H})$  the induced Lie algebra representation of  $\mathfrak{g}$ . The **moment map** associated to the representation  $\Phi$  is the function

$$\Omega : \mathbb{P}\mathcal{H} \rightarrow i\mathfrak{k}^*, \quad (3.2.32)$$

which assigns to each ray  $\psi \in \mathbb{P}\mathcal{H}$  the linear function  $\Omega(\psi) : i\mathfrak{k} \rightarrow \mathbb{R}$  defined by

$$\forall X \in i\mathfrak{k}, \quad \Omega(\psi)(X) = \text{Tr}(P_\psi \phi(X)) = \frac{\langle v, \phi(X)v \rangle}{\langle v, v \rangle}, \quad (3.2.33)$$

where  $v \in \psi_\times$  is any non-zero representative vector in  $\psi$ .

In this way, the moment map  $\Omega : \mathbb{P}\mathcal{H} \rightarrow i\mathfrak{k}^*$  can be interpreted as a kind of non-commutative gradient of the Kempf-Ness function. Evidently, the minimum value of the Kempf-Ness function, which is related to the capacity of the vector  $v$ , and the gradient of the Kempf-Ness function, which is captured by the moment map evaluated on the ray containing  $v$ , are intimately related. The following result, known as the Kempf-Ness theorem, solidifies this connection and follows directly from observations made in [Remark 3.2.29](#).

**Theorem 3.2.31** (Kempf-Ness). *Let  $G = K_{\mathbb{C}}$  be the complexification of a compact connected Lie group  $K$ , let  $\Phi : G \rightarrow \text{GL}(\mathcal{H})$  be a rational representation of  $G$  on a complex finite-dimensional Hilbert space  $\mathcal{H}$  with  $K$ -invariant inner product, and let  $\Omega : \mathbb{P}\mathcal{H} \rightarrow i\mathfrak{k}^*$  be the moment map associated to the representation  $\Phi : G \rightarrow \text{GL}(\mathcal{H})$ . Then, for all non-zero vectors  $v \in \mathcal{H}_\times$ ,*

$$\forall g \in G : \|\Phi(g)v\| \geq \|v\| \iff \Omega([v]) = 0. \quad (3.2.34)$$

*Equivalently,  $\text{cap}(v) = \|v\|$  if and only if  $\Omega([v]) = 0$ .*

*Additionally, if  $\Omega([v]) = 0$  and  $w \in \Phi(G)v$  has the same norm as  $v$  ( $\|w\| = \|v\|$ ), then  $v$  and  $w$  necessarily belong to the same  $K$ -orbit,  $w \in \Phi(K)v$ .*

**Proposition 3.2.32.** *Let  $r \in \mathbb{N}$  and let  $K = \text{U}(1)^{\times r}$  be the  $r$ -dimensional torus. The complexification of  $K$  is therefore  $G = \mathbb{C}_\times^r$ . Let  $\Phi : G \rightarrow \text{GL}(\mathcal{V})$  be a representation of  $G$  where  $\mathcal{V}$  admits of the decomposition  $\mathcal{V} = \bigoplus_{\lambda \in \Lambda} \mathcal{V}_\lambda$  where  $\Lambda = \mathbb{Z}^r$  be a finite subset  $r$ -tuples of integers and where  $\mathcal{V}_\lambda$  is the multiplicity space for the irreducible representation of  $\mathbb{C}_\times^r$  with weight  $\lambda$ . Let  $v \in \mathcal{V}$  be a vector with weight-space decomposition  $v = \sum_\lambda v_\lambda$  (where  $v_\lambda \in \mathcal{V}_\lambda$ ) and define the support of  $v$  to be*

$$\Lambda_v := \{\lambda \in \mathbb{Z}^r \mid v_\lambda \neq 0\}. \quad (3.2.35)$$

*Then the capacity of the vector  $v$  vanishes if and only if the convex hull of  $\Lambda_v \subset \mathbb{Z}^r$ , viewed as a subset of  $\mathbb{R}^r$ , excludes zero.*

*Proof.* In this setting, the Lie algebra of  $K = \text{U}(1)^{\times r}$  is  $\mathfrak{k} = (i\mathbb{R})^r$  and thus  $i\mathfrak{k} = \mathbb{R}^r$ . Now a vector  $v \in \mathcal{V}$  has weight  $\lambda \in \mathbb{Z}^r$  if for all  $(z_1, \dots, z_r) \in \mathbb{C}_\times^r$ ,

$$\Phi(z_1, \dots, z_r)v = z_1^{\lambda_1} \cdots z_r^{\lambda_r} v. \quad (3.2.36)$$

The subspace consisting of all vectors in  $\mathcal{V}$  with weight  $\lambda$  is the (possibly empty) weight space  $\mathcal{V}_\lambda$ . Using the decomposition of  $v$  into distinct weight spaces, the capacity squared of  $v$  can be expressed as

$$\text{cap}^2(v) = \inf_{z \in \mathbb{C}_\times^r} \|\Phi(z)v\|^2 = \inf_{z \in \mathbb{C}_\times^r} \sum_{\lambda \in \Lambda_v} \|z_1^{\lambda_1} \cdots z_r^{\lambda_r} v_\lambda\|^2 = \inf_{x \in \mathbb{R}^r} \sum_{\lambda \in \Lambda_v} e^{\lambda \cdot x} \|v_\lambda\|^2. \quad (3.2.37)$$

where  $x = (x_1, \dots, x_r) \in \mathbb{R}^r$  has  $i$ th component  $x_i = 2 \log |z_i|$  and where

$$\lambda \cdot x = \langle \lambda, x \rangle_{\mathbb{R}^r} = \sum_{i=1}^r w_i x_i = \sum_{i=1}^r 2w_i \log |z_i|. \quad (3.2.38)$$

Notice that the summand,  $e^{\lambda \cdot x} \|v_\lambda\|^2$ , in Eq. (3.2.37) is strictly positive,  $\|v_\lambda\|^2 > 0$ , and therefore  $\text{cap}(v) = 0$  if and only if the value of  $\lambda \cdot x$  can be made arbitrarily negative by some  $x \in \mathbb{R}^r$  simultaneously for all  $\lambda \in \Lambda_v$ . In other words,  $\text{cap}(v) = 0$  if for all  $L \in \mathbb{R}_{\geq 0}$ , there exists an  $x \in \mathbb{R}^r$  such that  $\lambda \cdot x < -L$  holds for  $\lambda \in \Lambda_v$ . This condition, in turn, can be seen to be equivalent to the condition that the convex hull of  $\Lambda_v$  in  $\mathbb{R}^r$  does *not* contain the origin. This is because if there exists a  $c \in [0, 1]^{|\Omega(v)|}$  such that  $\sum_{\lambda \in \Lambda_v} c_\lambda \lambda = 0 \in \mathbb{Z}^r$  and  $\sum_{\lambda \in \Lambda_v} c_\lambda = 1$  (that is, the origin is in the convex hull of  $\Lambda_v$ ), then one concludes for any  $x \in \mathbb{R}^r$ , that  $\lambda \cdot x \geq 0$  holds for some  $\lambda \in \Lambda_v$  because

$$0 = \sum_{\lambda \in \Lambda_v} (c_\lambda \lambda) \cdot x = \sum_{\lambda \in \Lambda_v} c_\lambda (\lambda \cdot x) \quad (3.2.39)$$

On the other hand, if there exists a hyperplane  $h \in \mathbb{R}^r$  separating  $\Lambda_v$  from the origin, i.e.  $\lambda \cdot h > 0$  holds for all  $\lambda \in \Lambda_v$ , then taking  $x = -sh$  for  $s > 0$  arbitrarily large yields for all  $\lambda \in \Lambda_v$ ,

$$\lim_{s \rightarrow \infty} \lambda \cdot x = \lim_{s \rightarrow \infty} -s \lambda \cdot h = -\infty. \quad (3.2.40)$$

In summary, we have  $\text{cap}(v) = 0$  if and only if the convex hull of  $\Lambda_v \subset \mathbb{Z}^r$  (viewed as a subset of  $\mathbb{R}^r$ ) excludes zero.  $\square$

**Remark 3.2.33.** Regarding the moment map of the representation of the  $r$ -dimensional torus,  $U(1)^{\times r}$ , defined by Eq. (3.2.36), one has for all vectors  $v \in \mathcal{V}$  and associated rays  $[v] \in \mathbb{P}\mathcal{V}$  and  $x \in i\mathfrak{k} = \mathbb{R}^r$  the following relationship:

$$\Omega([v])(x) = \sum_{\lambda \in \Lambda_v} \frac{\|v_\lambda\|^2}{\|v\|^2} \lambda \cdot x. \quad (3.2.41)$$

In other words, the moment map evaluated on the ray  $[v]$ ,  $\Omega([v])$ , can be identified with an element in the convex hull of the weights  $\Lambda_v$  supporting  $v$ . Moreover, the particular convex weighting which identifies  $\Omega([v])$  is given by the coefficients  $\|v_\lambda\|^2 / \|v\|^2$ .

**Remark 3.2.34.** Another foundational result in geometric invariant theory is that the image of the moment map  $\Omega : \mathbb{P}\mathcal{H} \rightarrow i\mathfrak{k}^*$  evaluated on the orbit closure  $\overline{\Phi(G)v}$  always intersects the closed fundamental Weyl chamber,  $i\mathfrak{t}_+^*$ , (viewed as a subset of  $i\mathfrak{k}^*$  by using an inner product on  $i\mathfrak{k}$  which is invariant under the adjoint action  $\text{Ad}$  of  $K$  on  $i\mathfrak{k}$ ) forms a convex polytope known as the *moment polytope* of  $v$ , denoted by  $\Delta_v := \Omega(\overline{\Phi(G)v}) \cap i\mathfrak{t}_+^*$ . In Section 5.1.3, we will return to the topic of moment polytopes.

The purpose of the remainder of this section is to highlight the various compositional aspects and symmetries of moment maps and capacities associated to a representation  $\Phi : G \rightarrow \text{GL}(\mathcal{H})$  of a complex reductive group  $G = K_{\mathbb{C}}$  with maximal compact subgroup  $K$ . To begin, notice that as the capacity of a vector,  $v \in \mathcal{H}$ , is defined as an optimization over all of  $G$ , the action of  $G$  on  $v$  does not modify the capacity of  $v$ , i.e.,  $\text{cap}(\Phi(g)v) = \text{cap}(v)$  for all  $g \in G$ . Unlike the invariance of the capacity map, the moment map,  $\Omega([v])$ , of the ray containing  $v$  does vary with the action of  $G$ .



**Lemma 3.2.35.** *The moment map  $\Omega : \mathbb{P}\mathcal{H} \rightarrow i\mathfrak{k}^*$  associated to the representation  $\Phi : G \rightarrow \text{GL}(\mathcal{H})$  (where  $G = K_{\mathbb{C}}$  and  $K$  is a compact Lie group) is  $K$ -equivariant, meaning for all  $\psi \in \mathbb{P}\mathcal{H}$ , and  $k \in K$  the moment map satisfies*

$$\text{Ad}^*(k)(\Omega(\psi)) = \Omega(\Phi(k) \cdot \psi) \quad (3.2.42)$$

where  $\text{Ad}^* : G \rightarrow \text{GL}(\mathfrak{g}^*)$  is the dual of the adjoint representation.

*Proof.* Recall from [Section 2.4](#) that the adjoint representation  $\text{Ad} : G \rightarrow \text{GL}(\mathfrak{g})$  satisfies  $\text{Ad}(g)(X) = gXg^{-1}$  and is furthermore represented by  $\phi(\text{Ad}(g)(X)) = \phi(gXg^{-1}) = \Phi(g)\phi(X)\Phi(g^{-1})$ . Since  $k \in K$  is represented as a unitary,  $\Phi(k^{-1}) = \Phi(k)^*$ , we have for generic  $X \in i\mathfrak{k}$ , the following chain of equivalences

$$\begin{aligned} \text{Ad}^*(k)(\Omega(\psi))(X) &= \Omega(\psi)(\text{Ad}(k^{-1})(X)), \\ &= \text{Tr}(P_{\psi}\phi(\text{Ad}(k^{-1})(X))), \\ &= \text{Tr}(P_{\psi}\phi(k^{-1}Xk)), \\ &= \text{Tr}(P_{\psi}\Phi(k^{-1})\phi(X)\Phi(k)), \\ &= \text{Tr}(P_{\Phi(k)\cdot\psi}\phi(X)), \\ &= \Omega(\Phi(k) \cdot \psi)(X). \end{aligned} \quad (3.2.43)$$

□

### 3.2.4 Composition of capacities & moment maps

**Remark 3.2.36.** The following collection of results are concerned with the moment map or capacity map of tensor-products of multiple distinct representations, as defined in [Section 2.4.3](#). In order to distinguish between the moment maps or capacity map associated to distinct representations, a subscript will be added wherever appropriate to avoid confusion. Specifically, the moment map and capacity map associated to the representation  $\Phi : G \rightarrow \text{GL}(\mathcal{H})$  will be respectively written as

$$\begin{aligned} \Omega_{\Phi} : \mathbb{P}\mathcal{H} &\rightarrow i\mathfrak{k}^*, \\ \text{cap}_{\Phi} : \mathcal{H} &\rightarrow \mathbb{R}_{\geq 0} \end{aligned} \quad (3.2.44)$$

In summary, it will be shown that moment maps are additive along tensor products while capacity maps are supermultiplicative.

**Lemma 3.2.37.** *Let  $\Phi_1 : G \rightarrow \text{GL}(\mathcal{V}_1)$  and  $\Phi_2 : G \rightarrow \text{GL}(\mathcal{V}_2)$  be representations and let  $v_1 \otimes v_2 \in \mathcal{V}_1 \otimes \mathcal{V}_2$ . Then*

$$\text{cap}_{\Phi_1 \otimes \Phi_2}(v_1 \otimes v_2) \geq \text{cap}_{\Phi_1}(v_1)\text{cap}_{\Phi_2}(v_2) = \text{cap}_{\Phi_1 \boxtimes \Phi_2}(v_1 \otimes v_2). \quad (3.2.45)$$

*Proof.* The proof follows from the injectivity of the copying map  $\Delta : G \rightarrow G \times G$  sending  $g$  to  $\Delta(g) = (g, g)$  which connects the internal and external tensor products of  $\Phi_1$  and  $\Phi_2$  first encountered in [Remark 2.4.20](#).

$$\text{cap}_{\Phi_1 \otimes \Phi_2}(v_1 \otimes v_2) = \inf_{g \in G} \|(\Phi_1(g) \otimes \Phi_2(g))(v_1 \otimes v_2)\| \quad (3.2.46)$$

$$\geq \inf_{\substack{g_1 \in G \\ g_2 \in G}} \|\Phi_1(g_1)v_1\| \|\Phi_2(g_2)v_2\| = \text{cap}_{\Phi_1}(v_1)\text{cap}_{\Phi_2}(v_2). \quad (3.2.47)$$

□



**Lemma 3.2.38.** *Let  $\Phi_1 : G_1 \rightarrow \text{GL}(\mathcal{V}_1)$  and  $\Phi_2 : G_2 \rightarrow \text{GL}(\mathcal{V}_2)$  be representations and let  $v_1 \otimes v_2 \in \mathcal{V}_1 \otimes \mathcal{V}_2$ . Then the moment map for the external tensor product representation  $\Phi_1 \boxtimes \Phi_2$  is of the form  $\Omega_{\Phi_1 \boxtimes \Phi_2} : \mathbb{P}(\mathcal{V}_1 \otimes \mathcal{V}_2) \rightarrow (i\mathfrak{k}_1 \oplus i\mathfrak{k}_2)^*$  where*

$$\Omega_{\Phi_1 \boxtimes \Phi_2}([v_1 \otimes v_2]) = \Omega_{\Phi_1}([v_1]) \oplus \Omega_{\Phi_2}([v_2]). \quad (3.2.48)$$

*Proof.* If  $\mathfrak{g}_1$  and  $\mathfrak{g}_2$  are the Lie algebras of  $G_1$  and  $G_2$ , then the Lie algebra of  $G_1 \times G_2$  is  $\mathfrak{g}_1 \oplus \mathfrak{g}_2$ . Moreover, if  $\phi_1$  and  $\phi_2$  are the respective induced Lie algebra representations, then  $\Phi_1 \boxtimes \Phi_2$  induces the Lie algebra representation  $\phi_1 \boxtimes \phi_2 : \mathfrak{g}_1 \oplus \mathfrak{g}_2 \rightarrow \mathfrak{gl}(\mathcal{V}_1 \otimes \mathcal{V}_2)$  which for  $X_1 \oplus X_2 \in \mathfrak{g}_1 \oplus \mathfrak{g}_2$  is defined by

$$(\phi_1 \boxtimes \phi_2)(X_1 \oplus X_2) = \phi_1(X_1) \otimes I_{\mathcal{V}_2} + I_{\mathcal{V}_1} \otimes \phi_2(X_2). \quad (3.2.49)$$

Applying this result to the definition of the moment map in [Definition 3.2.30](#) when  $v = v_1 \otimes v_2$  yields the claim.  $\square$

**Lemma 3.2.39.** *Let  $\Phi_1 : G \rightarrow \text{GL}(\mathcal{V}_1)$  and  $\Phi_2 : G \rightarrow \text{GL}(\mathcal{V}_2)$  be representations and let  $v_1 \otimes v_2 \in \mathcal{V}_1 \otimes \mathcal{V}_2$ . Then the moment map for the internal tensor product representation  $\Phi_1 \otimes \Phi_2 : G \rightarrow \text{GL}(\mathcal{V}_1 \otimes \mathcal{V}_2)$  is of the form  $\Omega_{\Phi_1 \otimes \Phi_2} : \mathbb{P}(\mathcal{V}_1 \otimes \mathcal{V}_2) \rightarrow i\mathfrak{k}^*$  where*

$$\Omega_{\Phi_1 \otimes \Phi_2}([v_1 \otimes v_2]) = \Omega_{\Phi_1}([v_1]) + \Omega_{\Phi_2}([v_2]). \quad (3.2.50)$$

*Proof.* If  $\mathfrak{g}$  is the Lie algebra of  $G$ , then the representation of  $\mathfrak{g}$  induced by the internal tensor product representation  $\Phi_1 \otimes \Phi_2$  of  $G$  on  $\mathcal{V}_1 \otimes \mathcal{V}_2$  is simply  $\phi_1 \otimes \phi_2 : \mathfrak{g} \rightarrow \mathfrak{gl}(\mathcal{V}_1 \otimes \mathcal{V}_2)$  defined for  $X \in \mathfrak{g}$  by

$$(\phi_1 \otimes \phi_2)(X) = \phi_1(X) \otimes I_{\mathcal{V}_2} + I_{\mathcal{V}_1} \otimes \phi_2(X). \quad (3.2.51)$$

Applying this to the definition of the moment map in [Definition 3.2.30](#) when  $v = v_1 \otimes v_2$  yields the claim.  $\square$

## 3.3 Occasionality & semistability

### 3.3.1 Cumulants of quantum observables

This section considers the moment generating function associated to the random variable formed by a quantum state and quantum observable pair. The main result, [Corollary 3.3.2](#), is the derivation of an error-bound on the second order expansion of this moment generating function.

**Lemma 3.3.1.** *Let  $X$  be a self-adjoint operator,  $X^* = X$ , acting on a complex finite-dimensional Hilbert space  $\mathcal{H}$ . Let  $\text{End}(\mathcal{H})$  be a  $C^*$ -algebra of linear maps on  $\mathcal{H}$  and let  $\varphi : \text{End}(\mathcal{H}) \rightarrow \mathbb{C}$  be a state. Let  $R$  be the discrete random variable with distribution  $\text{Prob}(R = x) = \varphi(P_x)$  where  $P_x \in \text{End}(\mathcal{H})$  projects onto the eigenspace of  $X$  with eigenvalue  $x \in \mathbb{R}$  and let  $M : \mathbb{R} \rightarrow \mathbb{R}$  be its moment generating function:*

$$M(t) = \mathbb{E}(\exp(tR)) = \sum_x e^{tx} \varphi(P_x) = \varphi(\exp(tX)) \quad (3.3.1)$$

Whenever  $t \in \mathbb{R}$  is such that  $\varphi(\exp(tX)) \neq 0$  let  $\varphi_t$  denote the state sending any  $Y \in \text{End}(\mathcal{H})$  to

$$\varphi_t(Y) := \frac{\varphi(e^{t\frac{X}{2}} Y e^{t\frac{X}{2}})}{\varphi(e^{tX})}. \quad (3.3.2)$$

Then the first three derivatives of the cumulant generating function  $K(t) = \log M(t) = \log \varphi(\exp(tX))$  are:

$$K^{(1)} = \varphi_t(X), \quad (3.3.3)$$

$$K^{(2)} = \varphi_t(X^2) - \varphi_t(X)^2 = \varphi_t((X - \varphi_t(X))^2), \quad (3.3.4)$$

$$K^{(3)} = \varphi_t(X^3) - 3\varphi_t(X^2)\varphi_t(X) + 2\varphi_t(X)^3 = \varphi_t((X - \varphi_t(X))^3). \quad (3.3.5)$$

$$(3.3.6)$$

*Proof.* Throughout this proof, we make use of the assumption that  $M(t) = \varphi(e^{tX}) \neq 0$  for all  $t$  in a neighborhood of zero so that  $K(t) = \log M(t)$  is well-defined.

$$K^{(1)}(t) = \frac{\varphi(X e^{tX})}{\varphi(e^{tX})} \quad (3.3.7)$$

$$K^{(2)}(t) = \frac{\varphi(X^2 e^{tX})\varphi(e^{tX}) - \varphi(X e^{tX})^2}{\varphi(e^{tX})^2} \quad (3.3.8)$$

$$K^{(3)}(t) = \frac{(\varphi(X^3 e^{tX})\varphi(e^{tX}) - \varphi(X^2 e^{tX})\varphi(X e^{tX}))\varphi(e^{tX})^2}{\varphi(e^{tX})^4} - \frac{2(\varphi(X^2 e^{tX})\varphi(e^{tX}) - \varphi(X e^{tX})^2)\varphi(e^{tX})\varphi(X e^{tX})}{\varphi(e^{tX})^4} \quad (3.3.9)$$

$$= \frac{\varphi(X^3 e^{tX})\varphi(e^{tX})^2 - 3\varphi(X^2 e^{tX})\varphi(X e^{tX})\varphi(e^{tX}) + 2\varphi(X e^{tX})^3}{\varphi(e^{tX})^3} \quad (3.3.10)$$

The claim follows from noting that  $X^p e^{tX} = e^{t\frac{X}{2}} X^p e^{t\frac{X}{2}}$  holds for any exponent  $p \in \mathbb{N}$  because  $[e^{tX}, X] = 0$  and therefore

$$\frac{\varphi(X^p e^{tX})}{\varphi(e^{tX})} = \varphi_t(X^p). \quad (3.3.11)$$

□

**Corollary 3.3.2.** *Let everything be as defined by Lemma 3.3.1. The following inequalities hold for any  $t \in \mathbb{R}$  where  $\varphi(e^{tX}) \neq 0$  so that  $K(t)$  is well-defined.*

$$0 \leq |K^{(3)}(t)| \leq 2 \|X\|_{\text{op}} K^{(2)}(t), \quad 0 \leq K^{(2)}(t) \leq 4 \|X\|_{\text{op}}^2. \quad (3.3.12)$$

*Proof.* Let  $S_t = X - \varphi_t(X)$  be the self-adjoint operator,  $S_t^* = S_t$ , obtained by shifting  $X$  about its mean  $\varphi_t(X)$ . By subadditivity of the operator norm, we have

$$\|S_t\|_{\text{op}} \leq \|X\|_{\text{op}} + \varphi_t(X) \|1\|_{\text{op}} \leq 2 \|X\|_{\text{op}}. \quad (3.3.13)$$

Furthermore, the first few moments of  $S_t$  with respect to  $\varphi_t$  are

$$\varphi_t(S_t) = 0, \quad \varphi_t(S_t^2) = K^{(2)}, \quad \varphi_t(S_t^3) = K^{(3)}. \quad (3.3.14)$$

Then by Eq. (2.2.23) and positivity of  $\varphi_t$ ,

$$0 \leq \varphi_t(S_t^2) \leq \|S_t^2\|_{\text{op}} \leq 4 \|X\|_{\text{op}}^2 \quad (3.3.15)$$

Furthermore by Corollary 3.3.2,

$$|\varphi_t(S_t^3)| \leq \varphi_t(S_t^2)^{\frac{3}{2}} \leq 2 \|X\|_{\text{op}} K^{(2)}. \quad (3.3.16)$$

□

**Remark 3.3.3.** Before stating and proving the next result, let  $\epsilon > 0$  be small and let  $X \in \text{End}(\mathcal{H})$  be self-adjoint. For any state,  $\varphi : \text{End}(\mathcal{H}) \rightarrow \mathbb{C}$ , one can make the following approximation:

$$\varphi(\exp(\epsilon X)) \approx 1 + \epsilon \varphi(X) + \frac{\epsilon^2}{2} \varphi(X^2). \quad (3.3.17)$$

If additionally  $\varphi(X) = 0$ , one obtains the approximation  $\varphi(\exp(\epsilon X)) \approx 1 + \frac{\epsilon^2}{2} \varphi(X^2)$  and therefore up to terms of order  $\epsilon^3$ , we have

$$\log \varphi(\exp(\epsilon X)) \approx \frac{1}{2} \epsilon^2 \varphi(X^2). \quad (3.3.18)$$

The following result quantifies the error introduced by this approximation.

**Lemma 3.3.4.** *Let  $X^* = X$  be a self-adjoint operator and let  $\varphi$  be a state. Then the moment generating function satisfies*

$$\varphi(\exp(tX)) = \exp \left[ \varphi(X)t + (\varphi(X^2) - \varphi(X)^2) \frac{t^2}{2} + r(t) \right] \quad (3.3.19)$$

where  $r : \mathbb{R} \rightarrow \mathbb{C}$  is a remainder term that satisfies for all  $t \in \mathbb{R}$ ,

$$|r(t)| \leq \frac{4}{3} \|X\|_{\text{op}}^3 t^3. \quad (3.3.20)$$

Moreover when  $\varphi(X) = 0$ ,

$$\lim_{n \rightarrow \infty} \varphi \left( \exp \left( \frac{X}{\sqrt{n}} \right) \right)^n = \exp \left[ \frac{\varphi(X^2)}{2} \right] \quad (3.3.21)$$

*Proof.* The proof relies on taking a Taylor series of the cumulant generating function  $K(t) = \log \varphi(e^{tX})$  for  $t > 0$  about  $t = 0$  to second order

$$K(t) = \kappa_1 t + \frac{\kappa_2}{2!} t^2 + r(t), \quad (3.3.22)$$

where (i) the coefficient  $\kappa_n$  is the cumulant of degree  $n$ , i.e., the  $n$ th derivative of  $K(t)$  evaluated at  $t = 0$ , and (ii) the remainder term  $r(t)$  is (for a fixed  $t > 0$ ) of the form

$$r(t) = \frac{K^{(3)}(c)}{3!} t^3. \quad (3.3.23)$$

for some  $c \in [0, t]$ . Furthermore using [Corollary 3.3.2](#), the remainder term may be bounded uniformly with respect to  $t$  by

$$|r(t)| \leq \frac{1}{6} |K^{(3)}(c)| t^3 \leq \frac{4}{3} \|X\|_{\text{op}}^3 t^3. \quad (3.3.24)$$

Now using [Lemma 3.3.1](#) together with the fact that  $\varphi_t$  evaluated at  $t = 0$  is merely  $\varphi$  yields

$$\kappa_1 = \varphi(X), \quad \kappa_2 = \varphi((X - \varphi(X))^2). \quad (3.3.25)$$

Therefore if  $\varphi(X) = 0$ ,  $\kappa_1 = 0$  and  $\kappa_2 = \varphi(X^2)$  which means  $K(t)$  is to leading order quadratic in  $t$ :

$$K(t) = \varphi(X^2) \frac{t^2}{2} + r(t). \quad (3.3.26)$$

By appropriately scaling the cumulant generating function

$$K(t) \mapsto \lambda^2 K(\lambda^{-1}t) \quad (3.3.27)$$

for some large  $\lambda > 0$ , one can suppress the contribution of the remainder term while leaving the second order term unaffected in the sense that

$$\lambda^2 K(\lambda^{-1}t) = \varphi(X^2) \frac{t^2}{2} + \lambda^2 r(\lambda^{-1}t), \quad (3.3.28)$$

where by [Eq. \(3.3.24\)](#),

$$|\lambda^2 r(\lambda^{-1}t)| \leq \lambda^{-1} \frac{4}{3} \|H\|_{\text{op}}^3 t^3. \quad (3.3.29)$$

Setting  $t = 1$  and taking the limit as  $\lambda \rightarrow \infty$  produces

$$\lim_{\lambda \rightarrow \infty} \lambda^2 K(\lambda^{-1}) = \frac{\varphi(X^2)}{2} \quad (3.3.30)$$

as claimed. □

### 3.3.2 Typical, occasional & exceptional behaviours

In the forthcoming sections, namely [Section 3.3.3](#) and [Section 3.3.4](#), we will consider the asymptotics of sequences of probabilities  $\{p_n \in [0, 1] \mid n \in \mathbb{N}\}$ . In particular, we will be interested in the sequence of probabilities that arises from applying a quantum state  $\varphi : \text{End}(\mathcal{H}) \rightarrow \mathbb{C}$  to the sequence of fixed subspaces ([Section 3.2.2](#)) of increasing degree  $n$ ,  $\{(\mathcal{H}^{\otimes n})^\Phi \mid n \in \mathbb{N}\}$  associated to the representation  $\Phi : G \rightarrow \text{GL}(\mathcal{H})$  of a group  $G$ . Specifically, the sequence of probabilities will have the form

$$p_n = \varphi^{\otimes n}(\Gamma_{\Phi^{\otimes n}}) \quad (3.3.31)$$

where  $\Gamma_{\Phi^{\otimes n}} \in \text{End}(\mathcal{H}^{\otimes n})$  is the projection operator on the fixed subspace  $(\mathcal{H}^{\otimes n})^\Phi$ . It will be shown that the qualitative behaviour of this sequence of probabilities in the limit of large  $n$  encodes information about the relationship between the state  $\varphi$  and the group  $G$ . The purpose of this section is to define three types of asymptotic behaviours which we refer to as *typical*, *occasional* and *exceptional*.

**Definition 3.3.5.** A sequence  $\{p_n \in [0, 1] \mid n \in \mathbb{N}\}$  of probability assignments is said to describe

i) a *typical* behaviour if

$$\lim_{n \rightarrow \infty} p_n = 1, \quad (3.3.32)$$

ii) an *occasional* behaviour if there exists constants  $\alpha > 0, \beta > 0$  such that

$$\limsup_{n \rightarrow \infty} n^\alpha p_n \geq \beta, \quad (3.3.33)$$

iii) an *exceptional* behaviour if there exists a constant  $\gamma \in [0, 1)$  such that

$$\limsup_{n \rightarrow \infty} p_n^{\frac{1}{n}} \leq \gamma. \quad (3.3.34)$$

**Remark 3.3.6.** A *typical* behaviour is essentially any sequence of events with probabilities  $\{p_n \in [0, 1] \mid n \in \mathbb{N}\}$  that one can be arbitrarily certain will eventually occur because for any small  $\epsilon > 0$ , we have  $p_n > 1 - \epsilon$  for sufficiently large  $N$ .

**Remark 3.3.7.** While the notion of an *occasional* behaviour is more subtle, it faithfully captures the intuitive idea that a fair coin occasionally yields an equal number of heads and tails. Indeed, the probability of flipping a fair coin  $n$  times and obtaining an equal number of heads and tails is equal to

$$p_n = \begin{cases} \frac{1}{2^n} \binom{n}{\frac{n}{2}} & n \text{ even} \\ 0 & n \text{ odd.} \end{cases} \quad (3.3.35)$$

Although obtaining an equal number of heads and tails is *atypical* in the sense that  $p_n \rightarrow 0$  as  $n \rightarrow \infty$ , one expects to obtain an equal number of heads and tails *occasionally* because one can show (for  $n$  even) that  $p_n \geq (2n)^{-1/2}$  and therefore [Eq. \(3.3.33\)](#) holds for  $\alpha = \frac{1}{2}$  and  $\beta = \frac{1}{\sqrt{2}}$ .

**Remark 3.3.8.** The terminology of an *exceptional* behaviour is justified because as  $n \rightarrow \infty$ , the probability  $p_n$  decays to zero exponentially fast. To see this, let  $\gamma \in [0, 1)$  be as in [Eq. \(3.3.34\)](#) and let  $\epsilon > 0$  be such that  $\gamma + \epsilon < 1$ . Then for sufficiently large  $N \in \mathbb{N}$ , the limit in [Eq. \(3.3.34\)](#) implies  $\sup_{n \geq N} p_n^{\frac{1}{n}} \leq \gamma + \epsilon$ . This implies that  $p_n \leq (\gamma + \epsilon)^n$  which means  $p_n$  eventually decays to zero at a rate that is at least exponential in  $n$  with exponent  $\log(\gamma + \epsilon) < 0$ . In other words, there exists an  $r \in (0, 1)$  such that  $p_n \leq r^n$  holds for sufficiently large  $n \geq N$ .

When the sequence of probabilities under investigation arises from the application of a tensor power state  $\varphi^{\otimes n}$  to the projection operator  $\Gamma_{\Phi^{\otimes n}}$  of a fixed subspace  $(\mathcal{H}^{\otimes n})^\Phi \subset \mathcal{H}^{\otimes n}$  as in [Eq. \(3.3.31\)](#), one can use the operator inequality  $\Gamma_{\Phi^{\otimes n}} \otimes \Gamma_{\Phi^{\otimes m}} \leq \Gamma_{\Phi^{\otimes(n+m)}}$  (see [Remark 3.2.16](#)) to show that the sequence  $\{p_n \mid n \in \mathbb{N}\}$  is *super-multiplicative* in the sense that

$$p_{n+m} = \varphi^{\otimes(n+m)}(\Gamma_{\Phi^{\otimes(n+m)}}) \geq \varphi^{\otimes(n+m)}(\Gamma_{\Phi^{\otimes n}} \otimes \Gamma_{\Phi^{\otimes m}}) = \varphi^{\otimes n}(\Gamma_{\Phi^{\otimes n}}) \varphi^{\otimes m}(\Gamma_{\Phi^{\otimes m}}) = p_n p_m. \quad (3.3.36)$$

As it turns out, this property is useful for establishing limits such as those appearing in Eq. (3.3.34). First we recall a powerful lemma for subadditive sequences of real numbers known as Fekete's subadditivity lemma [Ste97, Lem. 1.2.1].

**Lemma 3.3.9.** *For every sequence  $\{a_n\}_{n \in \mathbb{N}}$  satisfying*

$$a_{n+m} \leq a_n + a_m, \quad (3.3.37)$$

*the following limit exists and satisfies*

$$\lim_{n \rightarrow \infty} \frac{a_n}{n} = \inf_{n \in \mathbb{N}} \frac{a_n}{n}. \quad (3.3.38)$$

*Proof.* If the sequence ever reaches  $a_m = -\infty$  for finite  $m$ , then by subadditivity, both sides of the above equation are  $-\infty$ . Henceforth assume  $a_n > -\infty$  for all  $n$ . Now for any  $n = mk + r$ , subadditivity implies  $a_n \leq ma_k + a_r$  and thus

$$\frac{a_n}{n} \leq \frac{ma_k}{n} + \frac{a_r}{n} = \left(1 - \frac{r}{n}\right) \frac{a_k}{k} + \frac{a_r}{n}. \quad (3.3.39)$$

Moreover for fixed  $k$ , one may assume for all  $n \geq k$  that  $n = km + r$  holds for some  $r$  satisfying  $0 \leq r < k$ . For any  $\epsilon > 0$  pick a  $k$  such that  $\frac{a_k}{k}$  is within  $\epsilon$  of  $L = \inf_{n \in \mathbb{N}} \frac{a_n}{n}$ :

$$\frac{a_k}{k} < L + \epsilon. \quad (3.3.40)$$

Since  $k$  is now fixed and  $r < k$ ,

$$\lim_{n \rightarrow \infty} \frac{a_n}{n} < \lim_{n \rightarrow \infty} \left[ \left(1 - \frac{r}{n}\right) (L + \epsilon) + \frac{a_r}{n} \right] = L + \epsilon. \quad (3.3.41)$$

Since this holds for any  $\epsilon > 0$ , the lemma holds by taking  $\epsilon \rightarrow 0$ .  $\square$

**Remark 3.3.10.** Note that every super-multiplicative sequence  $\{b_n\}_{n \in \mathbb{N}}$ , meaning  $b_{n+m} \geq b_n b_m$ , gives rise to a subadditive sequence with  $a_n = -\log b_n$  (assuming  $b_n > 0$  for all  $n$ ). This implies an analogue of Fekete's lemma for super-multiplicative sequences.

**Corollary 3.3.11.** *For every sequence  $\{b_n\}_{n \in \mathbb{N}}$  satisfying*

$$b_{n+m} \geq b_n b_m, \quad (3.3.42)$$

*the following limit exists and satisfies*

$$\limsup_{n \rightarrow \infty} b_n^{\frac{1}{n}} = \sup_{n \in \mathbb{N}} b_n^{\frac{1}{n}}. \quad (3.3.43)$$

*Proof.* By supermultiplicativity, if  $b_k > 0$  for some  $k$ , then  $b_{mk} \geq b_k^m > 0$  for all  $m$ . Therefore  $a_m \mapsto -\log b_{mk}$  is a well-defined subadditive sequence which implies  $\lim_{m \rightarrow \infty} \frac{a_m}{m} = \inf_m \frac{a_m}{m}$ . Therefore for all  $k$  such that  $b_k > 0$ ,

$$\lim_{m \rightarrow \infty} b_{mk}^{\frac{1}{mk}} = \sup_{m \in \mathbb{N}} b_{mk}^{\frac{1}{mk}} \geq b_k^{\frac{1}{k}}. \quad (3.3.44)$$

This implies that

$$\limsup_{n \rightarrow \infty} b_n^{\frac{1}{n}} \geq \sup_{n \in \mathbb{N}} b_n^{\frac{1}{n}}, \quad (3.3.45)$$

and the reverse inequality holds trivially.  $\square$

**Remark 3.3.12.** Note that one cannot necessarily replace  $\limsup$  with  $\lim$  in the above corollary because it remains possible to find a subsequence  $\{n_j\}_{j \in \mathbb{N}}$  contained outside of all sub-semigroups  $\{mk | k \in \mathbb{N}\}$  of  $\mathbb{N}$  such that  $b_{n_j} = 0$  holds for all  $j$ . See the footnote in [FW20] for further explanation of this subtlety.

### 3.3.3 The occasionality theorem

The purpose of this section is to develop and prove an important relationship between the constant subspaces and semistable vectors of a representation  $\Phi : G \rightarrow \mathrm{GL}(\mathcal{H})$  and its tensor powers  $\Phi^{\otimes n} : G \rightarrow \mathrm{GL}(\mathcal{H})$  (Theorem 3.3.16). The proof presented here is taken directly from Franks and Walter [FW20]. Our only deviation from [FW20] is to emphasize the role of cumulants (Section 2.1.3), as noticed by [Bür+19, Rem. 3.16]. As this result serves as the foundation for many other results in this thesis, the proof is represented here purely for completeness.

Before doing so, it will be necessary to distinguish between subgroups of  $G$  which stabilize a given non-zero vector  $v \in \mathcal{H}_x$  from the subgroups of  $G$  which stabilize a given ray  $\psi \in \mathbb{P}\mathcal{H}$ . To clarify this subtle distinction, we consider the following example.

**Example 3.3.13.** Let  $\mathcal{H}_+$  and  $\mathcal{H}_-$  be two orthogonal and complementary subspaces of a complex finite-dimensional Hilbert space

$$\mathcal{H} = \mathcal{H}_+ \oplus \mathcal{H}_-, \quad (3.3.46)$$

and let  $P_{\pm}$  be the orthogonal projection operator onto  $\mathcal{H}_{\pm}$ . Moreover, let  $X = P_+ - P_- \in \mathrm{End}(\mathcal{H})$  be the operator which reflects vectors through the subspace  $\mathcal{H}_+ \subset \mathcal{H}$ , i.e.,  $X = \mathbb{1}_{\mathcal{H}} - 2P_-$ . Then consider any vector  $v \in \mathcal{H}_x$  with non-zero component in both  $\mathcal{H}_+$  and  $\mathcal{H}_-$ . Let  $\Phi : \mathbb{C}_x \rightarrow \mathrm{GL}(\mathcal{H})$  the representation of  $\mathbb{C}_x$  defined by  $\Phi(e^z) = \exp(zX)$ . The action of  $\mathbb{C}_x$  on  $v = v_+ + v_-$  thus

$$\Phi(e^z)v = e^{+z}v_+ + e^{-z}v_-, \quad (3.3.47)$$

and therefore the stabilizer of  $v$  is the trivial subgroup  $\{1\} \subset \mathbb{C}_x$  while the stabilizer of the ray containing  $v$  is the subgroup  $\mathbb{Z}_2 \simeq \{-1, +1\} \subset \mathbb{C}_x$ .

The following result, found in [MN84, Lem. 2.2], demonstrates the connection between geometric and algebraic notions of stability.

**Lemma 3.3.14.** *Let  $\psi \in \mathbb{P}\mathcal{H}$  be a ray and let  $v \in \psi_x$  be non-zero so that  $\mathbb{C}v = \psi$ . Let  $\Phi : G \rightarrow \mathrm{GL}(\mathcal{H})$  be a representation of a reductive group  $G$  and let  $G_{\psi}$  denote the subgroup of  $G$  which leaves the subspace  $\psi$  invariant and  $G_v$  the subgroup of  $G_{\psi}$  which leaves the vector  $v$  invariant. Then either  $G_{\psi}/G_v \simeq \mathbb{C}_x$  or  $G_{\psi}/G_v$  is finite (and thus  $\dim G_{\psi} = \dim G_v$ ). Moreover if  $G_{\psi}/G_v \simeq \mathbb{C}_x$  then  $v$  is necessarily unstable.*

*Proof.* The quotient  $G_{\psi}/G_v$  of  $G_{\psi}$  by  $G_v$  considered by the lemma arises when considering the sequence

$$G_v \longrightarrow G_{\psi} \longrightarrow \mathbb{C}_x \quad (3.3.48)$$

where the map  $G_v \rightarrow G_{\psi}$  is simply a subgroup inclusion (whose image is the subgroup  $G_v \subseteq G_{\psi}$ ) while  $G_{\psi} \rightarrow \mathbb{C}_x$  is the map sending  $g \in G_{\psi}$  to the unique non-zero complex

number  $c_g$  satisfying  $\Phi(g)v = c_g v$ . Evidently, the kernel of the latter map  $G_\psi \rightarrow \mathbb{C}_\times$ , i.e., those  $g \in G_\psi$  such that  $\Phi(g)v = v$  is equal to  $G_v$ , and thus the above sequence is an example of an *exact* sequence. In any case, the quotient  $G_\psi/G_v$  is precisely the image of the action of  $G_\psi$  on  $v$  and is necessarily an algebraic subgroup of  $\mathbb{C}_\times$  (as  $G$  itself is reductive). Being an algebraic subgroup of  $\mathbb{C}_\times$ ,  $G_\psi/G_v$  must either be isomorphism to i) the entire group  $\mathbb{C}_\times$  or ii) a finite subgroup of  $\mathbb{C}_\times$ . If  $G_\psi/G_v \simeq \mathbb{C}_\times$ , then there exists an element  $X$  in the Lie algebra of  $G_\psi$  such that  $\Phi(\exp(tX))v = e^{tx}v$  for some non-zero  $x \in \mathbb{R}_\times$ . Therefore, in the limit of either large or small  $t \in \mathbb{R}$ ,  $\Phi(\exp(tX))v = e^{tx}v \rightarrow 0$ , and thus the vector  $v$  must be unstable.  $\square$

A very useful corollary of [Lemma 3.3.14](#), which applies to semistable vectors, is the following result.

**Corollary 3.3.15.** *Let  $\psi \in \mathbb{P}\mathcal{H}$  and let  $v \in \psi_\times$  be a unit vector that is semistable with respect to a reductive group representation  $\Phi : G \rightarrow \mathrm{GL}(\mathcal{H})$ . Then there exists a positive integer  $m \in \mathbb{N}$  such that for all unitary group elements  $u \in G$  (meaning  $u^* = u^{-1}$ ),*

$$\langle v, \Phi(u)v \rangle^m = \mathrm{Tr}(P_\psi \Phi(u))^m = 1 \iff u \in G_\psi. \quad (3.3.49)$$

*Proof.* By [Lemma 3.3.14](#) and the semistability of  $v$ ,  $G_\psi/G_v$  is a finite subgroup of  $\mathbb{C}_\times$ . Let the order of this finite abelian group be  $m \in \mathbb{N}$ . The reverse direction ( $\iff$ ) holds even if  $u$  is not unitary because if  $g \in G_\psi$ , then  $gG_v \in G_\psi/G_v$  and therefore  $g^m G_v = eG_v \simeq G_v$  is the identity coset because the order of  $gG_v$  must divide  $m$ . Therefore,

$$\begin{aligned} 1 &= \mathrm{Tr}(P_\psi \Phi(e)) = \mathrm{Tr}(P_\psi \Phi(eG_v)), \\ &= \mathrm{Tr}(P_\psi \Phi(g^m G_v)) = \mathrm{Tr}(P_\psi \Phi(g)^m), \\ &= \mathrm{Tr}(P_\psi \Phi(g))^m \end{aligned} \quad (3.3.50)$$

where the last equality holds because  $g \in G_\psi$ . The forward direction ( $\implies$ ) holds because  $\mathrm{Tr}(P_\psi \Phi(u))^m = 1$  implies  $\mathrm{Tr}(P_\psi \Phi(u))$  is an  $m$ th root of unity and thus a element of  $U(1)$ . Therefore  $\Phi(u)v$  (which is necessarily a unit vector by unitarity of  $u$ ) is colinear with  $v$  and thus  $u$  stabilizes  $\mathbb{C}v = \psi$ .  $\square$

The main result of this section, which we are now in a position to prove, may be found in [\[FW20, Prop. 3.1\]](#).

**Theorem 3.3.16.** *Let  $\Phi : G \rightarrow \mathrm{GL}(\mathcal{H})$  and let  $\psi \in \mathbb{P}\mathcal{H}$  have vanishing moment map with respect to  $\Phi$ , i.e.  $\Omega(\psi) = 0$ . Then there exists positive integers  $m, c \in \mathbb{N}$  such that*

$$\lim_{k \rightarrow \infty} (mk)^{c/2} \mathrm{Tr}(P_\psi^{\otimes mk} \Gamma_{\Phi^{\otimes mk}}) = \sqrt{\frac{(2\pi)^c}{\det(Q_\psi)}} > 0 \quad (3.3.51)$$

where  $\Gamma_{\Phi^{\otimes n}}$  is the projection operator onto the fixed subspace of degree  $n$  ([Definition 3.2.15](#)), and  $Q_\psi$  is the quadratic form defined by  $X \mapsto \mathrm{Tr}(\phi(X)P_\psi\phi(X)^*)$ .



*Proof.* Our proof follows almost exactly the proof in [FW20]. Begin by letting  $v \in \mathcal{H}$  be any unit vector contained in the ray  $\psi \in \mathbb{P}\mathcal{H}$ . Then the pure state  $P_\psi$  is of the form  $\text{Tr}(P_\psi X) = \langle v, Xv \rangle$ . Applying Eq. (3.2.15) produces

$$\text{Tr}(P_\psi^{\otimes n} \Gamma_{\Phi^{\otimes n}}) = \int_{u \in K} du \text{Tr}(P_\psi \Phi(u))^n = \int_{u \in K} du \langle v, \Phi(u)v \rangle^n. \quad (3.3.52)$$

Now consider the function  $u \mapsto \text{Tr}(P_\psi \Phi(u)) = \langle v, \Phi(u)v \rangle$  appearing in the integrand above. As every  $u \in K$  is represented by a unitary matrix  $\Phi(u)$ , the Cauchy-Schwarz inequality yields the following bounds on the magnitude of  $\text{Tr}(P_\psi \Phi(u))$ :

$$0 \leq |\text{Tr}(P_\psi \Phi(u))| \leq 1. \quad (3.3.53)$$

The primary idea of this proof is to note that when  $u \in K$  satisfies  $|\text{Tr}(P_\psi \Phi(u))| < 1$  and  $n \rightarrow \infty$ , the integrand,  $\text{Tr}(P_\psi \Phi(u))^n$ , converges to zero:

$$|\text{Tr}(P_\psi \Phi(u))| < 1 \implies \lim_{n \rightarrow \infty} \text{Tr}(P_\psi \Phi(u))^n = 0. \quad (3.3.54)$$

On the other hand, when  $|\text{Tr}(P_\psi \Phi(u))| = 1$  one concludes that  $\text{Tr}(P_\psi \Phi(u)) = e^{i\theta}$  holds for some  $\theta \in [0, 2\pi]$  that is implicitly dependent on  $u$ . Fortunately, by the Kempf-Ness theorem (Theorem 3.2.31) the vanishing of the moment map for  $\psi$  implies  $v \in \psi_\times$  is semistable which, by Corollary 3.3.15, implies the existence of a positive integer  $m$  (equal to the order of the finite group  $K_\psi/K_v \subset \text{U}(1)$ ) such that  $\text{Tr}(P_\psi \Phi(u))^{mk} = 1$  holds if and only if  $u$  belongs to the identity coset  $eK_\psi \in K/K_\psi$ . This observation enables one to partially integrate Eq. (3.3.52) to obtain

$$\text{Tr}(P_\psi^{\otimes mk} \Gamma_{\Phi^{\otimes mk}}) = \int_{\bar{u} \in K/K_\psi} d\bar{u} \text{Tr}(P_\psi \Phi(\bar{u}))^{mk}, \quad (3.3.55)$$

where  $d\bar{u}$  is the unique normalized left- $K$ -invariant measure on the right cosets  $K/K_\psi$ .

By Eq. (3.3.54), the only portion of  $K/K_\psi$  which may contribute to the integral in the limit of large  $k$  is therefore those cosets,  $\bar{u} \in K/K_\psi$ , which are close to the identity. In order to define this collection of cosets, which will be denoted by  $U \subset K/K_\psi$  in Eq. (3.3.56) below, first let  $\mathfrak{k}_\psi$  denote the Lie algebra of the stabilizer  $K_\psi$  and identify the orthogonal complement of  $\mathfrak{k}_\psi^\perp$  of  $\mathfrak{k}_\psi \subseteq \mathfrak{k}$  (taken with respect to an inner product on  $\mathfrak{k}$  which is invariant under the adjoint representation of  $K$ ) with the tangent space of  $K/K_\psi$  at the identity coset using the projection map  $K \mapsto K/K_\psi$  and its surjective differential [FW20]. Let  $\epsilon > 0$  and let  $B_\epsilon(0) \subset \mathfrak{k}_\psi^\perp$  denote the open ball of radius  $\epsilon$  around the origin in  $\mathfrak{k}_\psi^\perp$ . Let  $\text{Exp} : \mathfrak{k}_\psi^\perp \rightarrow K/K_\psi$  denote the exponential map derived by first restricting the original exponential map  $\exp : \mathfrak{k} \rightarrow K$  for  $K$  to the subspace  $\mathfrak{k}_\psi^\perp$  and then second projecting through the quotient  $K \mapsto K/K_\psi$ . For  $\epsilon$  sufficiently small, the restriction of  $\text{Exp}$  to  $B_\epsilon(0)$  is a local diffeomorphism whose image is the open subset  $U \subset K/K_\psi$ , i.e.

$$U := \text{Exp}(B_\epsilon(0)) = \{\text{Exp}(X) \in K/K_\psi \mid X \in B_\epsilon(0) \subset \mathfrak{k}_\psi^\perp\}. \quad (3.3.56)$$

Let  $J : \mathfrak{k}_\psi^\perp \rightarrow \mathbb{R}$  denote the Jacobian of  $\text{Exp} : \mathfrak{k}_\psi^\perp \rightarrow K/K_\psi$ , which is a smooth function of  $X \in \mathfrak{k}_\psi^\perp$ . At the origin,  $0 \in \mathfrak{k}_\psi^\perp$ , the differential of  $\text{Exp}$  is an isometry and thus its determinant is

$J(0) = 1$  and therefore for sufficiently small  $\epsilon$ ,  $J(X) < 2$  holds for all  $A \in B_\epsilon(0)$ . Altogether, one obtains

$$\int_{\bar{u} \in U} d\bar{u} \text{Tr}(P_\psi \Phi(\bar{u}))^{mk} = \int_{A \in B_\epsilon(0)} dA J(A) \text{Tr}(P_\psi \exp(\phi(A)))^{mk}. \quad (3.3.57)$$

At this stage note that  $A \in \mathfrak{k}_\psi^\perp \subset \mathfrak{k}$  is necessarily a skew-Hermitian operator ( $A^* = -A$ ). Consequently, the integrand above, namely  $\text{Tr}(P_\psi \exp(\phi(A)))$ , is simply the characteristic function for the random variable with distribution  $\text{Tr}(P_\psi P_a)$  where  $P_a$  projects onto the eigenspace of  $A$  with eigenvalue  $a \in i\mathbb{R}$ . Since the moment map of  $\psi$  vanishes,  $\text{Tr}(P_\psi \phi(X)) = 0$  for all  $X \in i\mathfrak{k}$  and therefore [Lemma 3.3.4](#) becomes applicable. In order to apply [Lemma 3.3.4](#), one first needs to apply a change the variables  $A = \frac{A'}{\sqrt{mk}}$  and then take a limit as  $k \rightarrow \infty$ . As  $k \rightarrow \infty$ , we have  $J(\frac{A'}{\sqrt{mk}}) \rightarrow 1$  and  $B_{\sqrt{mk}\epsilon}(0) \rightarrow \mathfrak{k}_\psi^\perp$  so therefore,

$$\lim_{k \rightarrow \infty} (mk)^{\dim(\mathfrak{k}_\psi^\perp)/2} \int_{A \in B_\epsilon(0)} dA J(A) \text{Tr}(P_\psi \exp(\phi(A)))^{mk} \quad (3.3.58)$$

$$= \lim_{k \rightarrow \infty} \int_{A' \in B_{\sqrt{mk}\epsilon}(0)} dA' J\left(\frac{A'}{\sqrt{mk}}\right) \text{Tr}\left(P_\psi \exp\left(\phi\left(\frac{A'}{\sqrt{mk}}\right)\right)\right)^{mk}, \quad (3.3.59)$$

$$= \int_{A' \in \mathfrak{k}_\psi^\perp} dA' \lim_{k \rightarrow \infty} \text{Tr}\left(P_\psi \exp\left(\phi\left(\frac{A'}{\sqrt{mk}}\right)\right)\right)^{mk}. \quad (3.3.60)$$

The exchange between the integral and limit is permitted by Lebesgue's dominated convergence theorem because the integrands above are bounded in absolute value (as  $J(A) < 2$  for sufficiently small  $\epsilon$  and  $\exp(\phi(A)) \in U(\mathcal{H})$  is unitary). Now applying [Lemma 3.3.4](#),

$$\lim_{k \rightarrow \infty} \text{Tr}\left(P_\psi \exp\left(\phi\left(\frac{A'}{\sqrt{mk}}\right)\right)\right)^{mk} = \exp\left(-\frac{\text{Tr}(P_\psi \phi(A')^2)}{2}\right), \quad (3.3.61)$$

and therefore the integral over  $U \subset \mathfrak{k}_\psi^\perp$  is approximately Gaussian and thus

$$\begin{aligned} \lim_{k \rightarrow \infty} (mk)^{\dim(\mathfrak{k}_\psi^\perp)/2} \int_{\bar{u} \in U} d\bar{u} \text{Tr}(P_\psi \Phi(\bar{u}))^{mk} \\ = \int_{A' \in \mathfrak{k}_\psi^\perp} dA' \exp\left(-\frac{\text{Tr}(P_\psi \phi(A')^2)}{2}\right) = \sqrt{\frac{(2\Phi)^{\dim(\mathfrak{k}_\psi^\perp)}}{\det(Q_\psi)}}, \end{aligned} \quad (3.3.62)$$

where  $Q_\psi$  is the non-degenerate quadratic form sending  $X \in i\mathfrak{k}_\psi^\perp$  to

$$\text{Tr}(P_\psi \phi(X)^2) = \|\phi(X)v\|^2 > 0. \quad (3.3.63)$$

Since  $Q_\psi$  is positive definite, we have  $\det(Q_\psi) > 0$ .

As previously mentioned, the integral over the complement of  $U$  in  $\mathfrak{k}_\psi^\perp$  does not affect the above limit because the maximum  $C := \sup_{A \notin U} |\text{Tr}(P_\psi \exp(\phi(A)))|$  is obtained by some  $A^* \notin U$  (because the complement of  $U$  is compact and  $A \mapsto \text{Tr}(P_\psi \exp(\phi(A)))$  continuous) and therefore the integral

$$\int_{\bar{u} \notin U} d\bar{u} \text{Tr}(P_\psi \Phi(\bar{u}))^{mk} \leq 2 \int_{A \in B_\epsilon(0)} dA |\text{Tr}(P_\psi \exp(\phi(A)))|^{mk} \leq 2 \text{vol}(B_\epsilon(0)) C^{mk}, \quad (3.3.64)$$

converges to zero at an exponential rate with increasing  $k$  and therefore the polynomial prefactor  $(mk)^{\dim(\mathfrak{k}_\psi^\perp)/2}$  does not affect the limiting value.  $\square$

**Corollary 3.3.17.** *Let everything be as in [Theorem 3.3.16](#). Then there exists a constant,  $L_\Phi > 0$ , independent of  $\psi$ , such that for arbitrarily large  $n \in \mathbb{N}$ ,*

$$\mathrm{Tr}(P_\psi^{\otimes n} \Gamma_{\Phi^{\otimes n}}) \geq \frac{L_\Phi}{n^{\dim(\mathfrak{k})/2}}. \quad (3.3.65)$$

*Proof.* The proof uses the fact that the constant  $c = \dim(\mathfrak{k}_\psi^\perp)$  in the proof of [Theorem 3.3.16](#) can be bounded by

$$0 \leq \dim(\mathfrak{k}_\psi^\perp) \leq \dim(\mathfrak{k}). \quad (3.3.66)$$

Furthermore, since  $Q_\Phi(X) := \mathrm{Tr}(\phi(X)\phi(X)^*) \geq \mathrm{Tr}(\phi(X)P_\psi\phi(X)^*) = Q_\psi(X)$ , we have  $\det(Q_\Phi) \geq \det(Q_\psi)$ . Setting  $L_\Phi = \frac{1}{2}Q_\Phi^{-1/2}$  then proves the claim.  $\square$

The following corollary [Theorem 3.3.16](#) hides many of the constants appearing in [Theorem 3.3.16](#) which will be unnecessary when applied later in [Section 3.3.4](#).

**Corollary 3.3.18.** *Let  $\Phi : G \rightarrow \mathrm{GL}(\mathcal{H})$  and let  $\psi \in \mathbb{P}\mathcal{H}$  have vanishing moment map, i.e.  $\Omega(\psi) = 0$ . Then,*

$$\limsup_{n \rightarrow \infty} \mathrm{Tr}(P_\psi^{\otimes n} \Gamma_{\Phi^{\otimes n}})^{\frac{1}{n}} = 1 \quad (3.3.67)$$

where  $\Gamma_{\Phi^{\otimes n}}$  is the projection operator onto the fixed subspace of degree  $n$  ([Definition 3.2.15](#)).

*Proof.* Since the moment map of  $\psi$  vanishes, we can apply [Theorem 3.3.16](#) to obtain positive integers  $m, c \in \mathbb{N}$  such that the following limit, denoted by  $L$ , exists and is strictly positive.

$$L := \lim_{k \rightarrow \infty} (mk)^{c/2} \mathrm{Tr}(P_\psi^{\otimes mk} \Gamma_{\Phi^{\otimes mk}}) > 0. \quad (3.3.68)$$

Therefore, for any  $\epsilon \in (0, L)$ , there exists a sufficiently large  $K \in \mathbb{N}$ , such that for all  $k \geq K$ ,

$$(mk)^{c/2} \mathrm{Tr}(P_\psi^{\otimes mk} \Gamma_{\Phi^{\otimes mk}}) \geq L' := L - \epsilon > 0, \quad (3.3.69)$$

which implies for all  $k \geq K$ ,

$$1 \geq \mathrm{Tr}(P_\psi^{\otimes mk} \Gamma_{\Phi^{\otimes mk}}) \geq \frac{L'}{(mk)^{c/2}}. \quad (3.3.70)$$

Taking the  $(mk)$ -th root of both sides therefore yields for all  $k \geq K$ ,

$$1 \geq \mathrm{Tr}(P_\psi^{\otimes mk} \Gamma_{\Phi^{\otimes mk}})^{\frac{1}{mk}} \geq \left( \frac{L'}{(mk)^{c/2}} \right)^{\frac{1}{mk}}. \quad (3.3.71)$$

Therefore,

$$1 \geq \limsup_{n \rightarrow \infty} \mathrm{Tr}(P_\psi^{\otimes n} \Gamma_{\Phi^{\otimes n}})^{\frac{1}{n}} \geq \limsup_{n \rightarrow \infty} \left( \frac{L'}{n^{c/2}} \right)^{\frac{1}{n}} = 1, \quad (3.3.72)$$

which proves the claim.  $\square$

### 3.3.4 Strong duality

The following is [FW20, Lem. 2.2].

**Lemma 3.3.19.** *Let  $v \in \mathcal{V}$  and  $\Phi : G \rightarrow \mathrm{GL}(\mathcal{V})$  a representation of  $G$ . Then for all  $n \in \mathbb{N}$ ,*

$$\mathrm{cap}(v) \geq \left\| \Gamma_{\Phi^{\otimes n}} v^{\otimes n} \right\|^{\frac{1}{n}}. \quad (3.3.73)$$

*Proof.* For every  $n \in \mathbb{N}$  and  $g \in G$ ,

$$\|\Phi(g)v\| = \left\| \Phi(g)^{\otimes n} v^{\otimes n} \right\|^{\frac{1}{n}} \geq \left\| \Gamma_{\Phi^{\otimes n}} \Phi(g)^{\otimes n} v^{\otimes n} \right\|^{\frac{1}{n}} = \left\| \Gamma_{\Phi^{\otimes n}} v^{\otimes n} \right\|^{\frac{1}{n}}. \quad (3.3.74)$$

The claim follows because  $\mathrm{cap}(v)$  is defined as the largest lower bound on  $\|\Phi(g)v\|$ .  $\square$

Combining the lower bound on capacity offered by Lemma 3.3.19, together with the lower bound on the norm of  $\Gamma_{\Phi^{\otimes n}} v^{\otimes n}$  offered by Theorem 3.3.16 (or more accurately Corollary 3.3.18), one can prove [FW20, Thm. 1.1], which is stated below.

**Theorem 3.3.20.** *Let  $\Phi : G \rightarrow \mathrm{GL}(\mathcal{V})$  be a representation of a complex reductive group  $G$ . Then for all vectors  $v \in \mathcal{V}$ ,*

$$\mathrm{cap}(v) = \inf_{g \in G} \|\Phi(g)v\| = \limsup_{n \rightarrow \infty} \left\| \Gamma_{\Phi^{\otimes n}} v^{\otimes n} \right\|^{\frac{1}{n}}. \quad (3.3.75)$$

*Proof.* That  $\mathrm{cap}(v) \geq \limsup_{n \rightarrow \infty} \left\| \Gamma_{\Phi^{\otimes n}} v^{\otimes n} \right\|^{\frac{1}{n}}$  holds for any vector  $v$  follows directly from Lemma 3.3.19. The reverse inequality can be deduced by first recognizing that the map  $v \mapsto \left\| \Gamma_{\Phi^{\otimes n}} v^{\otimes n} \right\|^{\frac{1}{n}}$  is  $G$ -invariant and continuous (see the proof of Lemma 3.2.17) and thus constant on  $G$ -orbit closures. Moreover, if  $w \in \overline{\Phi(g)v}$  is in the  $G$ -orbit closure of  $v$ , the  $G$ -orbit closure of  $w$  satisfies  $\overline{\Phi(g)w} \subseteq \overline{\Phi(g)v}$  and thus  $\mathrm{cap}(v) \leq \mathrm{cap}(w)$ . In other words, to prove the reverse inequality, it suffices to prove  $\mathrm{cap}(w) = \limsup_{n \rightarrow \infty} \left\| \Gamma_{\Phi^{\otimes n}} w^{\otimes n} \right\|^{\frac{1}{n}}$  for some  $w \in \overline{\Phi(g)v}$ .

If  $v$  is unstable, then zero belongs to its orbit closure,  $0 \in \overline{\Phi(g)v}$ , and then Lemma 3.3.19 holds because the capacity of the zero vector is trivially zero,  $\mathrm{cap}(0) = 0$ . On the other hand, if  $v$  is semistable, then there exists a non-zero vector  $w \in \mathcal{V}_\times$  in the closure of the  $G$ -orbit of  $v$  such that i) the ray  $[w] \in \mathbb{P}\mathcal{V}$  has vanishing moment map, i.e.  $\Omega([w]) = 0$ , and ii) the vector  $w$  has norm attaining the capacity of  $v$ , i.e.  $\mathrm{cap}(v) = \mathrm{cap}(w) = \|w\|$ . Applying Corollary 3.3.18 to the ray  $[w] \in \mathbb{P}\mathcal{V}$  then yields

$$\limsup_{n \rightarrow \infty} \mathrm{Tr}(P_{[w]}^{\otimes n} \Gamma_{\Phi^{\otimes n}})^{\frac{1}{n}} = \limsup_{n \rightarrow \infty} \left( \frac{\left\| \Gamma_{\Phi^{\otimes n}} w^{\otimes n} \right\|^2}{\|w\|^2} \right)^{\frac{1}{n}} = 1, \quad (3.3.76)$$

Therefore, we have

$$\mathrm{cap}^2(w) = \|w\|^2 = \limsup_{n \rightarrow \infty} \left\| \Gamma_{\Phi^{\otimes n}} w^{\otimes n} \right\|^{\frac{2}{n}}. \quad (3.3.77)$$

$\square$

The following result, which provides the converse to Corollary 3.3.18, weakens the statement of Theorem 3.3.20 to consider the special case when the moment map vanishes and the capacity is maximized.

**Corollary 3.3.21.** *Let  $\Phi : G \rightarrow \mathrm{GL}(\mathcal{V})$  be a representation of a complex reductive group  $G$  with associated moment map  $\Omega : \mathbb{P}\mathcal{H} \rightarrow i\mathfrak{k}^*$ . Let  $\psi \in \mathbb{P}\mathcal{H}$ , then*

$$\Omega(\psi) = 0 \implies \limsup_{n \rightarrow \infty} \mathrm{Tr}(P_\psi^{\otimes n} \Gamma_{\Phi^{\otimes n}})^{\frac{1}{n}} = 1, \quad (3.3.78)$$

$$\Omega(\psi) \neq 0 \implies \limsup_{n \rightarrow \infty} \mathrm{Tr}(P_\psi^{\otimes n} \Gamma_{\Phi^{\otimes n}})^{\frac{1}{n}} < 1. \quad (3.3.79)$$

*Proof.* The result follows from i) the correspondence between a vanishing moment map and maximum capacity provided by the Kempf-Ness theorem [Theorem 3.2.31](#), ii) the strong-duality theorem, [Theorem 3.3.20](#), obtained above, and iii) the fact that

$$\mathrm{Tr}(P_\psi^{\otimes n} \Gamma_{\Phi^{\otimes n}}) = \frac{\|\Gamma_{\Phi^{\otimes n}} v\|^2}{\|v\|^2} \quad (3.3.80)$$

for any non-zero  $v \in \psi_\times \in \mathcal{H}$ . □

One way to interpret the statement of [Corollary 3.3.21](#) is as an asymptotic test for whether or not the moment map of a given ray vanishes. Specifically, [Corollary 3.3.21](#) demonstrates that  $\Omega(\psi) \neq 0$  if and only if the sequence of probabilities,  $\mathrm{Tr}(P_\psi^{\otimes n} \Gamma_{\Phi^{\otimes n}}) \in [0, 1]$ , approaches zero at an *exponential* rate with increasing  $n$ .

# Chapter 4

## Estimation theory

### 4.1 Estimating a probability distribution

Consider a classic problem of estimation theory, namely, the problem of estimating the probability distribution for a random process yielding a finite number of distinct outcomes. Suppose the number of outcomes is  $d \in \mathbb{N}$  and the outcomes are labelled by the elements of the set  $[d] = \{1, \dots, d\}$ . As is standard, the set of all probability distributions with  $d$  outcomes can be identified with the  $(d - 1)$ -dimensional simplex,  $\Delta_d$ , consisting of tuples of  $d$  non-negative real numbers summing to one:

$$\Delta_d := \{(p_1, \dots, p_d) \in \mathbb{R}_{\geq 0}^d \mid p_1 + \dots + p_d = 1\}. \quad (4.1.1)$$

Assume that the outcomes of the process are produced independently on repeated trials. Furthermore, assume that the individual outcomes are produced according a known distribution, given by

$$q = (q_1, \dots, q_d) \in \Delta_d. \quad (4.1.2)$$

Under these assumptions, the probability of producing a sequence of type  $\lambda = (\lambda_1, \dots, \lambda_d)$ , where  $\lambda_j \in \mathbb{N}$  denotes the occurrences of outcome  $j$ , is given by the multinomial distribution,

$$m_q(\lambda) = \frac{n!}{\lambda_1! \dots \lambda_d!} q_1^{\lambda_1} \dots q_d^{\lambda_d}. \quad (4.1.3)$$

where  $n = \lambda_1 + \dots + \lambda_d$  is the total number of outcomes. If instead the distribution  $q$  is not known, but a sequence of type  $\lambda$  has been observed, then a somewhat reasonable estimate for  $q$  is the distribution  $p$  which maximizes the value of  $m_p(\lambda)$ . A direct calculation reveals that the so-called maximum likelihood estimate,  $p$ , is equal to  $\lambda/n$ , i.e., the relative frequencies of each outcome,

$$p = (p_1, \dots, p_d) = \left( \frac{\lambda_1}{n}, \dots, \frac{\lambda_d}{n} \right). \quad (4.1.4)$$

In practice, when the length of the sequence is small, the maximum likelihood estimate  $p$  tends to be a poor estimate for the value of  $q$ . Of course, as  $n$  becomes larger, the quality of the estimate  $p$  gets better.

One particularly elegant way to quantify the relationship between source distribution  $q$  and the estimate  $p$  as a function of  $n$  is a classical result due to Sanov [San61; DZ10], which states

$$(n+1)^{-d} \exp(-nD(p\|q)) \leq m_q(\lambda) \leq \exp(-nD(p\|q)), \quad (4.1.5)$$

where  $D(p\|q)$  is a non-negative quantity known as the relative entropy or Kullback-Liebler divergence [Kul97; BF14], and is defined as

$$D(p\|q) = \sum_{j=1}^d p_j (\log p_j - \log q_j) \quad (4.1.6)$$

whenever  $q_j = 0$  implies  $p_j = 0$ , and  $D(p\|q) = \infty$  otherwise. Note that the relative entropy,  $D(p\|q)$ , only vanishes when  $p = q$ , and therefore two observations follow. First, the probability of producing an estimate  $p$  which deviates from the correct distribution  $q$  decays at an exponential rate with increasing  $n$ , as quantified by  $D(p\|q) > 0$ . Second, the probability of producing an estimate which is exactly equal to the correct distribution,  $p = q$ , and thus  $D(p\|q) = 0$ , can be lower bounded by the reciprocal of a factor that is polynomial in the value of  $n$ , i.e.,  $(n+1)^d$ .

In Chapter 3, we already witnessed a similar distinction between the exponential decay of probabilities on one hand (Lemma 3.3.19), and the polynomial lower bound on probabilities on the other (Theorem 3.3.16). The purpose of this chapter is expand upon this connection in the context of estimating properties of quantum processes. To begin, we wish to demonstrate how to recover Eq. (4.1.5) using the techniques from Chapter 3.

In quantum theory, the most straight-forward way to model a random process with  $d$  distinct outcomes is to consider a  $d$ -dimensional vector space  $\mathbb{C}^d$ , with standard orthonormal basis  $\{e_1, \dots, e_d\}$ , along with a fixed vector  $v \in \mathbb{C}^d$  with unit norm  $\|v\| = 1$ . For each  $j \in [d]$ , let  $P_j = e_j e_j^*$  denote the orthogonal projection operator onto the subspace spanned by  $e_j$ . Then, the coefficients in the decomposition of the vector  $v$  into the standard orthonormal basis generate a probability distribution, given by

$$q_v := (\|P_1 v\|^2, \dots, \|P_d v\|^2) = (|v_1|^2, \dots, |v_d|^2) \in \Delta_d. \quad (4.1.7)$$

Note that fixing an orthonormal basis is equivalent to fixing a representation,  $\Phi : \mathrm{U}(1)^d \rightarrow \mathrm{U}(d)$ , of the  $d$ -dimensional torus,  $\mathrm{U}(1)^d$ , acting on  $\mathbb{C}^d$ , such that the matrix form of the representation is equal to

$$\Phi(e^{i\theta_1}, \dots, e^{i\theta_d}) = \mathrm{diag}(e^{i\theta_1}, \dots, e^{i\theta_d}) = \begin{pmatrix} e^{i\theta_1} & 0 & \dots & 0 \\ 0 & e^{i\theta_2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & e^{i\theta_d} \end{pmatrix}. \quad (4.1.8)$$

Since  $\mathrm{U}(1)^d$  is an abelian group, all of its irreducible representations are one-dimensional; indeed, for each  $j \in [d]$ , the subspace spanned by  $e_j$  is an invariant subspace of the representation in Eq. (4.1.8). In general, the irreducible representations of  $\mathrm{U}(1)^d$  are indexed by highest weights, which for  $\mathrm{U}(1)^d$ , can be identified by  $d$ -tuples of non-negative integers  $\lambda = (\lambda_1, \dots, \lambda_d) \in \mathbb{N}_{\geq 0}^d$ , where  $\Phi_\lambda$  has the form

$$\Phi_\lambda(e^{i\theta_1}, \dots, e^{i\theta_d}) = e^{i(\lambda_1 \theta_1 + \dots + \lambda_d \theta_d)}. \quad (4.1.9)$$

Now consider the  $n$ th tensor power,  $\Phi^{\otimes n} : \mathrm{U}(1)^d \rightarrow \mathrm{U}(d^n)$ , of the representation from Eq. (4.1.8). The multiplicity of irreducible representation  $\Phi_\lambda$ , indexed by  $\lambda = (\lambda_1, \dots, \lambda_d)$ , appearing inside the representation  $\Phi^{\otimes n}$  then corresponds to the multinomial coefficient,

$$\mathrm{Tr}(\Pi_{\Phi^{\otimes n}}^\lambda) = \frac{n!}{\lambda_1! \cdots \lambda_d!}. \quad (4.1.10)$$

Furthermore, if  $\Pi_{\Phi^{\otimes n}}^\lambda$  denotes the projection operator onto the subspace of weight  $\lambda$ , then the probability distribution associated to the decomposition of  $v^{\otimes n}$  into the various weight spaces is equal to the multinomial distribution associated to distribution  $q_v$ ,

$$m_{q_v}(\lambda) = \|\Pi_{\Phi^{\otimes n}}^\lambda v^{\otimes n}\|^2 = \frac{n!}{\lambda_1! \cdots \lambda_d!} |v_1|^{2\lambda_1} \cdots |v_d|^{2\lambda_d}. \quad (4.1.11)$$

Our strategy for deriving the bounds on  $m_{q_v}(\lambda)$  from Eq. (4.1.5) is to consider the action of the complexified torus on  $v \in \mathbb{C}^d$ . Since the Lie algebra of  $K = \mathrm{U}(1)^d$  consists of  $d$  purely imaginary numbers  $\mathfrak{k} = (i\theta_1, \dots, i\theta_d) \in i\mathbb{R}^d$ , the complexified Lie algebra consists of all complex numbers,  $\mathfrak{k} \oplus i\mathfrak{k} = \mathbb{C}^d$ , and the corresponding complexified Lie group is the *complex* torus,  $G = \mathbb{C}_\times^d$ . The representation in Eq. (4.1.8) then lifts to a representation for  $\mathbb{C}_\times^d$  of the form

$$\Phi(e^{x_1+i\theta_1}, \dots, e^{x_d+i\theta_d}) = \mathrm{diag}(e^{x_1+i\theta_1}, \dots, e^{x_d+i\theta_d}). \quad (4.1.12)$$

Now consider an element  $x = (x_1, \dots, x_d) \in \mathbb{R}^d \cong i\mathfrak{k}$ , along with the group element  $g \in \mathbb{R}_{>0}^d \subseteq \mathbb{C}_\times^d$  with completely real exponents,

$$g = e^x := (e^{x_1}, \dots, e^{x_d}). \quad (4.1.13)$$

In this setting, the moment map evaluated on the ray  $[v] \in \mathbb{P}(\mathbb{C}^d)$  satisfies

$$\Omega([v])(x) = \langle v, \phi(x)v \rangle = \sum_{j=1}^d x_j |v_j|^2, \quad (4.1.14)$$

In other words, the moment map,  $\Omega([v])$ , is essentially equivalent to the probability distribution  $q_j = |v_j|^2$  defined by Eq. (4.1.7). Furthermore, when  $g = (e^{x_1}, \dots, e^{x_d})$  acts on  $v$ , we obtain a new vector,  $v_x \in \mathbb{C}^d$ , of the form

$$v_x := \Phi(e^x)v = (e^{x_1}v_1, \dots, e^{x_d}v_d) \quad (4.1.15)$$

with norm squared

$$\|v_x\|^2 = \|\Phi(e^x)v\|^2 = \sum_{j=1}^d e^{2x_j} |v_j|^2. \quad (4.1.16)$$

Using the fact that  $\Phi(g^{-1})\Phi(g) = \mathbb{1}$  for all  $g$ , one obtains

$$\|\Pi_{\Phi^{\otimes n}}^\lambda v^{\otimes n}\| = \|\Pi_{\Phi^{\otimes n}}^\lambda \Phi(e^{-x})^{\otimes n} \Phi(e^x)^{\otimes n} v^{\otimes n}\| = e^{-\langle \lambda, x \rangle} \|\Pi_{\Phi^{\otimes n}}^\lambda v_x^{\otimes n}\|, \quad (4.1.17)$$

where the prefactor  $e^{-\langle \lambda, x \rangle}$  arises from the fact that  $\Pi_{\Phi^{\otimes n}}^\lambda$  is a weight space of weight  $\lambda$  and thus  $\Pi_{\Phi^{\otimes n}}^\lambda \Phi(e^{-x})^{\otimes n} = e^{-\langle \lambda, x \rangle} \Pi_{\Phi^{\otimes n}}^\lambda$ . Letting  $u_x := v_x / \|v_x\|$  be the normalization of  $v_x$  (note that  $v = u_0$ ), Eq. (4.1.17) may be expressed as

$$\|\Pi_{\Phi^{\otimes n}}^\lambda v^{\otimes n}\| = e^{-\langle \lambda, x \rangle} \|v_x\|^n \|\Pi_{\Phi^{\otimes n}}^\lambda u_x^{\otimes n}\|. \quad (4.1.18)$$



If  $p = (p_1, \dots, p_d)$  is now defined according to Eq. (4.1.4), so that  $np = \lambda$ , we conclude

$$\left\| \Pi_{\Phi^{\otimes n}}^\lambda v^{\otimes n} \right\| = (e^{-\langle p, x \rangle} \|v_x\|)^n \left\| \Pi_{\Phi^{\otimes n}}^\lambda u_x^{\otimes n} \right\|. \quad (4.1.19)$$

From Eq. (4.1.19) we can derive bounds on the likelihood  $\left\| \Pi_{\Phi^{\otimes n}}^\lambda v^{\otimes n} \right\|^2$  from bounds on  $\left\| \Pi_{\Phi^{\otimes n}}^\lambda u_x^{\otimes n} \right\|^2$ . To obtain the upper bound in Eq. (4.1.5) it suffices to note that  $\Pi_{\Phi^{\otimes n}}^\lambda$  is a projection operator and  $u_x$  is a unit norm vector so that

$$\left\| \Pi_{\Phi^{\otimes n}}^\lambda u_x^{\otimes n} \right\| \leq 1. \quad (4.1.20)$$

Then, by optimizing over all  $x \in \mathbb{R}^d$ , we obtain the upper bound

$$\left\| \Pi_{\Phi^{\otimes n}}^\lambda v^{\otimes n} \right\|^2 \leq \exp(-nI_v(p)), \quad (4.1.21)$$

where

$$I_v(p) = -\log \inf_{x \in \mathbb{R}^d} e^{-2\langle p, x \rangle} \|v_x\|^2 = -\log \inf_{x \in \mathbb{R}^d} \left( e^{-2\langle p, x \rangle} \sum_j e^{2x_j} |v_j|^2 \right). \quad (4.1.22)$$

Note that the optimization problem in Eq. (4.1.22) has a trivial solution of  $I_v(p) = \infty$  whenever there exists an index  $j$  such that  $p_j > 0$  while  $v_j = 0$ . Otherwise, the minimum is attained when  $x = y \in \mathbb{R}^d$  where

$$e^{2y_j} = \frac{p_j}{|v_j|^2}, \quad (4.1.23)$$

in which case

$$I_v(p) = \sum_{j=1}^d p_j (\log p_j - \log |v_j|^2). \quad (4.1.24)$$

Indeed, the function  $I_v(p)$ , as defined by Eq. (4.1.22), is equal to the relative entropy previously defined by Eq. (4.1.6).

In order to derive the lower-bound in Eq. (4.1.5), it suffices to consider the case where  $I_v(p) < \infty$ , since otherwise the lower-bound in Eq. (4.1.5) becomes trivial. In this case, we have  $x = y$  and therefore  $\|v_y\|^2 = \sum_{j=1}^d p_j = 1$  which means the distribution associated to the unit vector,  $u_y = v_y$ , is simply  $p_j = |(u_y)_j|^2 = \lambda_j/n$ . In this case the prefactor in Eq. (4.1.19) equals

$$(e^{-\langle p, y \rangle} \|v_y\|)^n = \exp(-nI_v(p)). \quad (4.1.25)$$

Furthermore, as the ratio of  $|(u_y)_j|^2$  to  $\lambda_j$  is independent of  $n$ , we can also conclude that  $\lambda$  maximizes the function  $\mu \mapsto \left\| \Pi_{\Phi^{\otimes n}}^\mu u_y^{\otimes n} \right\|^2$  [DZ10]. This follows because Eq. (4.1.11) implies<sup>1</sup>

$$\frac{\left\| \Pi_{\Phi^{\otimes n}}^\lambda u_y^{\otimes n} \right\|^2}{\left\| \Pi_{\Phi^{\otimes n}}^\mu u_y^{\otimes n} \right\|^2} = \prod_{j=1}^d \frac{\mu_j!}{\lambda_j!} p_j^{\lambda_j - \mu_j} \geq \prod_{j=1}^d \left( \frac{p_j}{\lambda_j} \right)^{\lambda_j - \mu_j} = 1. \quad (4.1.26)$$

Since the number of distinct weights in the decomposition of  $\Phi^{\otimes n}$  is at most  $(n+1)^d$ , Eq. (4.1.26) implies the likelihood  $\left\| \Pi_{\Phi^{\otimes n}}^\lambda u_y^{\otimes n} \right\|$  admits of the lower-bound

$$\left\| \Pi_{\Phi^{\otimes n}}^\lambda u_y^{\otimes n} \right\| \geq \frac{1}{(n+1)^d}. \quad (4.1.27)$$

---

<sup>1</sup>Here we use the inequality  $n!/m! \geq m^{n-m}$  which holds for all  $n, m \in \mathbb{N}$ .

Altogether, Eq. (4.1.19) becomes

$$\|\Pi_{\Phi^{\otimes n}}^\lambda v^{\otimes n}\| = (e^{-\langle p, y \rangle} \|v_y\|)^n \|\Pi_{\Phi^{\otimes n}}^\lambda u_y^{\otimes n}\| \geq \frac{1}{(n+1)^d} \exp(-nI_v(p)). \quad (4.1.28)$$

The purpose of this chapter is to generalize the above technique to consider starting with representations of non-commutative groups. By doing so, we will come to view the estimation of quantum states and their properties as a natural generalization of the estimation of a probability distribution.

## 4.2 Quantum measurements & representations

In this section we review a number of familiar constructions of covariant positive operator valued measures from the perspective of representation theory [Chi+04; Hol11].

### 4.2.1 Irreducible measurements

The purpose of this section is to describe a variety of positive-operator valued measures which naturally arise in the context of irreducible representations of groups.

Next, we review the standard construction of a covariant POVM arising from an irreducible unitary representation.

**Lemma 4.2.1.** *Let  $\Phi : K \rightarrow \mathbf{U}(\mathcal{H})$  be an irreducible unitary representation of a compact group  $K$  on a  $d$ -dimensional Hilbert space  $\mathcal{H}$ . Let  $\mu : \Sigma(K) \rightarrow [0, 1]$  be the unique  $K$ -invariant normalized Haar measure on  $K$ . Then the function  $E_\psi : \Sigma(K) \rightarrow \mathcal{B}(\mathcal{H})$  defined for  $\Delta \in \Sigma(K)$  by*

$$E_\psi(\Delta) := d \int_{k \in \Delta} \Phi(k) P_\psi \Phi(k^{-1}) d\mu(k), \quad (4.2.1)$$

*is a positive-operator-valued measure over  $K$  acting on  $\mathcal{H}$ .*

*Proof.* The construction used here follows Example 2.2.20. That  $E_\psi(\Delta)$  is positive semidefinite for all  $\Delta \in \Sigma(K)$  follows from  $P_\psi$  being projective and thus positive semidefinite and  $\Phi(k)^* = \Phi(k^{-1})$  being unitary. That  $E_\psi$  is normalized follows from noting that  $E_\psi(K) : \mathcal{H} \rightarrow \mathcal{H}$  is a  $K$ -covariant operator acting on an irreducible representation space and therefore, by Schur's lemma, must be proportional to the identity on  $\mathcal{H}$ . As the integrand satisfies

$$\mathrm{Tr}(P_{k \cdot \psi}) = \mathrm{Tr}(\Phi(k) P_\psi \Phi(k^{-1})) = \mathrm{Tr}(P_\psi) = \dim(\psi) = 1, \quad (4.2.2)$$

and  $\mathrm{Tr}(\mathbb{1}_{\mathcal{H}}) = \dim(\mathcal{H}) = d$ , the prefactor of  $d = \dim(\mathcal{H})$  in the definition of  $p_\psi$  ensures that  $E_\psi(K) = \mathbb{1}_{\mathcal{H}}$ .  $\square$

**Example 4.2.2.** Let  $K = C_d$  be the cyclic group on  $d$  symbols, e.g.,  $[d] = \{1, \dots, d\}$ , generated by the permutation  $\pi$  sending  $i \in [d]$  to

$$\pi(i) = (i + 1) \bmod d. \quad (4.2.3)$$

The normalized Haar measure over  $C_d$  has density  $|C_d|^{-1} = d^{-1}$ . Let  $\Phi : C_d \rightarrow \mathbf{U}(d)$  be the unitary representation of  $C_d$  where  $\Phi(\pi)$  acts on the orthonormal basis  $\{e_1, \dots, e_d\}$  for

$\mathbb{C}^d$  by sending the basis vector  $e_i$  to  $\Phi(\pi)e_i = e_{\pi(i)}$ . For each index  $i$ , let  $\phi_i \in \mathbb{P}\mathbb{C}^d$  be the one-dimensional subspace spanned by  $e_i$ . Then for each positive integer  $m \in \mathbb{N}$  and each element  $\pi^m \in C_d$  in the cyclic group, the POVM defined in [Lemma 4.2.1](#) (assuming  $\psi := \phi_1$ ) satisfies

$$E_{\phi_1}(\{\pi^m\}) = P_{\phi_{(m+1) \bmod d}}, \quad (4.2.4)$$

and thus corresponds to the standard projective measurement associated to the aforementioned orthonormal basis.

When the ray  $\psi \in \mathcal{H}$  considered in [Lemma 4.2.1](#) exhibits non-trivial symmetries, the POVM  $E_\psi$  over the group  $K$  can be related to a POVM,  $\tilde{E}_\psi$ , over the *orbit* of  $\psi$  under the action of  $K$ .

**Remark 4.2.3.** Let  $\Phi_\lambda : K \rightarrow \text{U}(\mathcal{H}_\lambda)$  be an irreducible unitary representation of a compact connected Lie group  $K$  with highest weight  $\lambda$  relative to a fixed maximal torus  $T$  in  $K$ . Let  $v_\lambda \in \mathcal{H}_\lambda$  be a vector of highest weight and let  $\psi_\lambda := [v_\lambda] \in \mathbb{P}\mathcal{H}_\lambda$  be the highest weight ray. The orbit of the highest weight vector,  $v_\lambda \in \mathcal{H}_\lambda$ , is

$$K \cdot v_\lambda := \{k \cdot v_\lambda := \Phi_\lambda(k)v_\lambda \in \mathcal{H}_\lambda \mid k \in K\} \subseteq \mathcal{H}_\lambda. \quad (4.2.5)$$

The orbit of the highest weight ray,  $\psi_\lambda \in \mathbb{P}\mathcal{H}_\lambda$  containing  $v_\lambda$ , is<sup>2</sup>

$$K \cdot \psi_\lambda := \{k \cdot \psi_\lambda := [\Phi_\lambda(k)v_\lambda] \in \mathbb{P}\mathcal{H}_\lambda \mid k \in K\} \subseteq \mathbb{P}\mathcal{H}_\lambda. \quad (4.2.6)$$

Finally let  $\omega \in i\mathfrak{k}^*$  be generic. The group  $K$  acts on  $i\mathfrak{k}^*$  through the dual of the adjoint representation of  $K$  on  $\mathfrak{k}$  such that for all  $X \in i\mathfrak{k}$ ,

$$(k \cdot \omega)(X) := [\text{Ad}^*(k)(\omega)](X) = \omega(\text{Ad}(k^{-1})(X)) = \omega(k^{-1}Xk) \in \mathbb{R}. \quad (4.2.7)$$

The orbit of  $\omega \in i\mathfrak{k}^*$ , called the **coadjoint orbit** of  $\omega$ , is therefore

$$K \cdot \omega := \{k \cdot \omega := \text{Ad}^*(k)(\omega) \mid k \in K\} \subseteq i\mathfrak{k}^*. \quad (4.2.8)$$

**Example 4.2.4.** Let everything be defined as in [Lemma 4.2.1](#). Let  $K_\psi$  be the stabilizer subgroup of  $K$  with respect to  $\psi$ , i.e.,

$$K_\psi := \{k \in K \mid k \cdot \psi = \psi\}. \quad (4.2.9)$$

As the orbit  $K \cdot \psi$  of the ray  $\psi$  can be identified with the set of left-cosets  $K \backslash K_\psi$ , one can consider the function  $s_\psi : K \rightarrow K \cdot \psi \cong K \backslash K_\psi$  assigning each  $k \in K$  to the point  $k \cdot \psi$  in the orbit or  $kK_\psi \in K \backslash K_\psi$  in the coset space,

$$s_\psi(k) = k \cdot \psi \cong kK_\psi. \quad (4.2.10)$$

Then the pushforward of the POVM defined in [Lemma 4.2.1](#) through  $s_\psi$  is the POVM

$$\tilde{E}_\psi : \Sigma(K \cdot \psi) \rightarrow \mathcal{B}(\mathcal{H}), \quad (4.2.11)$$

---

<sup>2</sup>Rays belonging to the orbit of the highest weight ray correspond to generalized *coherent* states in the sense of Perelomov [[Per72](#)] (or more precisely, in the sense of Klyachko [[Kly02](#)]).

taking values in the orbit of  $K \cdot \psi$  of  $\psi$  in  $\mathbb{P}\mathcal{H}$ , and satisfies for all  $\Delta' \in \Sigma(K \cdot \psi)$ ,

$$\tilde{E}_\psi(\Delta') := d \int_{\phi \in \Delta'} d((s_\psi)_* \mu)(\phi) P_\phi = d \int_{k \in s_\psi^{-1}(\Delta')} d\mu(k) P_{k \cdot \psi}. \quad (4.2.12)$$

Moreover, as the orbit  $K \cdot \psi$  is a closed subset of the projective space  $\mathbb{P}\mathcal{H}$ , we can alternatively view  $\tilde{E}_\psi$  as a POVM over all of  $\mathbb{P}\mathcal{H}$  with support only on  $K \cdot \psi \subseteq \mathbb{P}\mathcal{H}$ .

**Example 4.2.5.** Suppose  $\Phi_\lambda : K \rightarrow \mathrm{U}(\mathcal{H}_\lambda)$  is an irreducible unitary representation of a compact connected Lie group  $K$  with highest weight  $\lambda$  and  $d_\lambda := \dim(\mathcal{H}_\lambda)$ . Let  $\psi_\lambda \in \mathbb{P}\mathcal{H}_\lambda$  be the highest weight ray. The standard POVM,  $E_{\psi_\lambda}$ , from [Lemma 4.2.1](#) will, in this case, be abbreviated simply by  $E_\lambda : \Sigma(K) \rightarrow \mathcal{B}(\mathcal{H}_\lambda)$ , where

$$dE_\lambda(k) = d_\lambda \Phi_\lambda(k) P_{\psi_\lambda} \Phi_\lambda(k^{-1}) d\mu(k). \quad (4.2.13)$$

Furthermore, if  $\Omega_\lambda : \mathbb{P}\mathcal{H}_\lambda \rightarrow i\mathfrak{k}^*$  is the moment map associated to  $\Phi_\lambda$ , we will consider the POVM

$$F_\lambda : \Sigma(i\mathfrak{k}^*) \rightarrow \mathcal{B}(\mathcal{H}_\lambda), \quad (4.2.14)$$

defined for  $\Delta' \in \Sigma(i\mathfrak{k}^*)$  as

$$F_\lambda(\Delta') = E_\lambda((\Omega_\lambda \circ s_{\psi_\lambda})^{-1}(\Delta')). \quad (4.2.15)$$

In other words,  $F_\lambda$  is the pushforward of  $E_\lambda$  through the map  $k \mapsto \Omega_\lambda(k \cdot \psi_\lambda)$ . If  $g : i\mathfrak{k}^* \rightarrow \mathbb{R}$  is a measurable function, then for all  $\rho \in \mathcal{S}(\mathcal{H}_\lambda)$ ,

$$\int_{\omega \in i\mathfrak{k}^*} g(\omega) \mathrm{Tr}(dF_\lambda(\omega) \rho) = d_\lambda \int_{k \in K} g(\Omega_\lambda(k \cdot \psi_\lambda)) \mathrm{Tr}(P_{k \cdot \psi_\lambda} \rho) d\mu(k) \quad (4.2.16)$$

where  $\Omega_\lambda(k \cdot \psi_\lambda) = \mathrm{Ad}^*(k)(\lambda)$ .

## 4.2.2 Completely reducible measurements

The construction of the covariant POVM provided in [Example 4.2.5](#) can be generalized to the case of a non-irreducible representation,  $\Phi : K \rightarrow \mathrm{U}(\mathcal{H})$ , provided the representation is completely reducible. As finite-dimensional representations of compact groups are completely reducible by [Theorem 2.4.25](#), one can apply the covariant POVM provided by [Lemma 4.2.1](#) to each of the irreducible components in the decomposition of  $\Phi$ . When the compact Lie group  $K$  is additionally connected, the components of this decomposition can be indexed by dominant, analytically integral highest weights by [Theorem 2.4.55](#). But first, it will be helpful to define some notation for describing the inclusion of an irreducible representation inside a completely reducible representation.

**Definition 4.2.6.** Let  $\Phi : K \rightarrow \mathrm{U}(\mathcal{H})$  be a unitary representation of a compact, connected Lie group  $K$  on a finite-dimensional complex Hilbert space  $\mathcal{H}$ . Let  $\mathfrak{k}$  be the Lie algebra of  $K$  and  $\mathfrak{t}$  the Lie algebra of a fixed maximal torus  $T$  in  $K$ . Furthermore, let the decomposition of  $\mathcal{H}$  into its irreducible invariant subspaces be given by

$$\mathcal{H} \cong \bigoplus_{\lambda \in \Lambda_+} \mathcal{H}_\lambda \otimes \mathcal{M}_\lambda^{\mathcal{H}} \quad (4.2.17)$$

where the sum is taken over dominant, analytically integral weights  $\Lambda_+ \subset (it)^*$ , where  $\mathcal{H}_\lambda$  supports an irreducible unitary representation,  $\Phi_\lambda : K \rightarrow \text{U}(\mathcal{H}_\lambda)$ , with highest weight  $\lambda$  and highest weight ray  $\psi_\lambda \in \mathcal{H}_\lambda$ , and where

$$\mathcal{M}_\lambda^\Phi := \text{Hom}_K(\mathcal{H}_\lambda, \mathcal{H}) \quad (4.2.18)$$

is the multiplicity space of isomorphic copies of  $\mathcal{H}_\lambda$  inside  $\mathcal{H}$  with dimension equal to the multiplicity of  $\lambda$ . The summand  $\mathcal{H}_\lambda \otimes \mathcal{M}_\lambda^\Phi$  in Eq. (4.2.17) above is referred to as the **isotypic subspace** for  $\lambda \in \Lambda_+$  in  $\mathcal{H}_\lambda$ .

**Definition 4.2.7.** Let everything be defined as in Definition 4.2.6. The linear map describing the inclusion of the isotypic subspace in  $\mathcal{H}$  will be denoted by

$$\iota_\Phi^\lambda : \mathcal{H}_\lambda \otimes \mathcal{M}_\lambda^\Phi \rightarrow \mathcal{H}. \quad (4.2.19)$$

The linear map  $\iota_\Phi^\lambda$  is: i) an isometry, meaning

$$(\iota_\Phi^\lambda)^*(\iota_\Phi^\lambda) = \mathbb{1}_{\mathcal{H}_\lambda} \otimes \mathbb{1}_{\mathcal{M}_\lambda^\Phi}, \quad (4.2.20)$$

and thus  $\tilde{\Pi}_\Phi^\lambda := (\iota_\Phi^\lambda)(\iota_\Phi^\lambda)^*$  is a projection operator onto the isotypic subspace, and ii)  $K$ -covariant, meaning for all  $k \in K$ ,

$$(\iota_\Phi^\lambda)^*\Phi(k)\iota_\Phi^\lambda = \Phi_\lambda(k) \otimes \mathbb{1}_{\mathcal{M}_\lambda^\Phi}. \quad (4.2.21)$$

The **multiplicity channel** for the highest weight  $\lambda$  in the representation  $\Phi$  is the quantum channel

$$\mathcal{I}_\Phi^\lambda : \mathcal{B}(\mathcal{H}_\lambda) \rightarrow \mathcal{B}(\mathcal{H}) \quad (4.2.22)$$

sending each operator  $X \in \mathcal{B}(\mathcal{H}_\lambda)$  to the operator  $\mathcal{I}_\Phi^\lambda(X) \in \mathcal{B}(\mathcal{H})$  defined by

$$\mathcal{I}_\Phi^\lambda[X] := (\iota_\Phi^\lambda)(X \otimes \mathbb{1}_{\mathcal{M}_\lambda^\Phi})(\iota_\Phi^\lambda)^*. \quad (4.2.23)$$

If  $\lambda$  is *not* a highest weight of the representation, then  $\dim(\mathcal{M}_\lambda^\Phi) = 0$  and therefore the channel  $\mathcal{I}_\Phi^\lambda$  is not well-defined. The multiplicity channel inherits the  $K$ -covariance property from  $\iota_\Phi^\lambda$  in the sense that

$$\Phi(k)\mathcal{I}_\Phi^\lambda[X]\Phi(k^{-1}) = \mathcal{I}_\Phi^\lambda[\Phi_\lambda(k)X\Phi_\lambda(k^{-1})]. \quad (4.2.24)$$

**Definition 4.2.8.** Let everything be defined as in Definition 4.2.6 and Definition 4.2.7. Let  $\psi_\lambda \in \mathcal{H}_\lambda$  be the unique highest weight ray for the irreducible representation  $\mathcal{H}_\lambda$ . The projection operator onto the **subspace of highest weight vectors** for the representation  $\Phi : K \rightarrow \text{U}(\mathcal{H})$  is defined as

$$\Pi_\Phi^\lambda = \mathcal{I}_\Phi^\lambda[P_{\psi_\lambda}] = (\iota_\Phi^\lambda)(P_{\psi_\lambda} \otimes \mathbb{1}_{\mathcal{M}_\lambda^\Phi})(\iota_\Phi^\lambda)^*. \quad (4.2.25)$$

Additionally, the projection operator onto the subspace of highest weight vectors rotated by the action of  $k \in K$  is

$$\Pi_\Phi^{\lambda,k} = \Phi(k)\Pi_\Phi^\lambda\Phi(k^{-1}) = \mathcal{I}_\Phi^\lambda[P_{k \cdot \psi_\lambda}]. \quad (4.2.26)$$

**Example 4.2.9.** For each highest weight,  $\lambda \in \Lambda_+$ , and irreducible representation  $\Phi_\lambda : K \rightarrow U(\mathcal{H}_\lambda)$  with dimension  $d_\lambda = \dim(\mathcal{H}_\lambda)$ , let the highest weight ray be  $\psi_\lambda \in \mathcal{H}_\lambda$ . Consider the  $\mathcal{B}(\mathcal{H}_\lambda)$ -valued POVM from [Example 4.2.5](#) of the form  $E_\lambda : \Sigma(K) \rightarrow \mathcal{B}(\mathcal{H}_\lambda)$ ,

$$dE_\lambda(k) = d_\lambda P_{k \cdot \psi_\lambda} d\mu(k). \quad (4.2.27)$$

Using the channel  $\mathcal{I}_\lambda^\Phi : \mathcal{B}(\mathcal{H}_\lambda) \rightarrow \mathcal{B}(\mathcal{H})$ , the POVM  $E_\lambda$  acting on  $\mathcal{H}_\lambda$  can be *lifted* to the POVM  $\mathcal{I}_\lambda^\Phi \circ E_\lambda$  on  $\mathcal{H}$ . By summing these lifted POVMs over all dominant, analytically integral elements  $\lambda \in \Lambda_+$ , one obtains a POVM over  $K \times \Lambda_+$  of the form

$$E_\Phi : \Sigma(K \times \Lambda_+) \rightarrow \mathcal{B}(\mathcal{H}) \quad (4.2.28)$$

where for  $\Delta_K \in \Sigma(K)$  and  $\Delta_{\Lambda_+} \in \Sigma(\Lambda_+)$ , we have

$$E_\Phi(\Delta_K \times \Delta_{\Lambda_+}) := \sum_{\lambda \in \Delta_{\Lambda_+}} \mathcal{I}_\Phi^\lambda[E_\lambda(\Delta_K)]. \quad (4.2.29)$$

Expanding everything out (including the channels  $\mathcal{I}_\Phi^\lambda$ ), we obtain the expression

$$E_\Phi(\Delta_K \times \Delta_{\Lambda_+}) = \sum_{\lambda \in \Delta_{\Lambda_+}} d_\lambda \int_{k \in \Delta_K} d\mu(k) \Pi_\Phi^{\lambda, k}, \quad (4.2.30)$$

Furthermore, the marginal of  $E_\Phi$  obtained by integrating over  $K$  simplifies to the projective measurement

$$E_\Phi(K \times \Delta_{\Lambda_+}) = \sum_{\lambda \in \Delta_{\Lambda_+}} \tilde{\Pi}_\Phi^\lambda, \quad (4.2.31)$$

where  $\tilde{\Pi}_\Phi^\lambda$  is projection operator onto the isotypic subspace  $\mathcal{H}_\lambda \otimes \mathcal{M}_\lambda^{\mathcal{H}}$ ,

$$\tilde{\Pi}_\Phi^\lambda = \mathcal{I}_\Phi^\lambda(\mathbb{1}_{\mathcal{H}_\lambda}) = (\iota_\Phi^\lambda)(\iota_\Phi^\lambda)^*. \quad (4.2.32)$$

Alternatively, if one instead considers the POVM from [Example 4.2.5](#) of the form  $F_\lambda : \Sigma(i\mathfrak{k}^*) \rightarrow \mathcal{B}(\mathcal{H}_\lambda)$ , summing over  $\lambda \in \Lambda_+$  yields the POVM

$$F_\Phi : \Sigma(i\mathfrak{k}^*) \rightarrow \mathcal{B}(\mathcal{H}), \quad (4.2.33)$$

which is defined for all  $\Delta \in \Sigma(i\mathfrak{k}^*)$  by

$$F_\Phi(\Delta) := \sum_{\lambda \in \Delta_{\Lambda_+}} \mathcal{I}_\Phi^\lambda[F_\lambda(\Delta)]. \quad (4.2.34)$$

In other words, for all measurable functions  $g : i\mathfrak{k}^* \rightarrow \mathbb{R}$  and  $\rho \in \mathcal{S}(\mathcal{H})$ ,  $F_\Phi$  has the form

$$\int_{\omega \in i\mathfrak{k}^*} g(\omega) \text{Tr}(dF_\Phi(\omega)\rho) = \sum_{\lambda \in \Lambda_+} d_\lambda \int_{k \in K} g(\Omega_\lambda(k \cdot \psi_\lambda)) \text{Tr}(\Pi_\Phi^{\lambda, k} \rho) d\mu. \quad (4.2.35)$$

### 4.3 Deformed strong duality

The purpose of this section is develop a variety of variations of the strong duality result from [Section 3.3.4](#) which will be useful in [Section 4.4.1](#).

### 4.3.1 The inversion trick

Before proceeding, recall from [Section 3.3.4](#), specifically [Corollary 3.3.21](#), that the large  $n$  asymptotics of the probability  $\text{Tr}(P_\psi^{\otimes n} \Gamma_{\Phi^{\otimes n}})$ , where  $\Gamma_{\Phi^{\otimes n}}$  is the projection operator onto the subspace of invariant-vectors in  $\mathcal{H}^{\otimes n}$ , is directly related to whether or not the moment map  $\Omega_\Phi(\psi)$  of  $\psi$  vanishes. Specifically, if  $\Omega_\Phi(\psi) \neq 0$ , then the probability  $\text{Tr}(P_\psi^{\otimes n} \Gamma_{\Phi^{\otimes n}})$  decays at an exponential rate with increasing  $n$ , and thus one expects the event associated to  $\Gamma_{\Phi^{\otimes n}}$  rarely occurs for large  $n$ , if ever. Otherwise, if  $\Omega_\Phi(\psi) = 0$ , then the probability  $\text{Tr}(P_\psi^{\otimes n} \Gamma_{\Phi^{\otimes n}})$  does *not* decay at an exponential rate; in fact, by [Theorem 3.3.16](#) one expects that the event associated to  $\Gamma_{\Phi^{\otimes n}}$  occasionally occurs (in the sense of [Section 3.3.2](#)).

If one is instead interested in the precise *value* of the moment map,  $\Omega_\Phi(\psi) \in i\mathfrak{k}^*$ , a natural question arises: does there exist a method, analogous to [Corollary 3.3.21](#), for testing whether or not the moment map equals a specific value, say  $\omega = \Omega_\Phi(\psi)$ ? The answer to this question, at least for certain values of  $x \in i\mathfrak{k}^*$ , is fortunately yes.

Roughly speaking, the way to generalize [Corollary 3.3.21](#) is to first find or construct a *known* quantum state,  $\psi' \in \mathbb{P}\mathcal{W}$ , belonging to a different Hilbert space  $\mathcal{W}$  along with a different representation  $\Phi' : G \rightarrow \text{GL}(\mathcal{W})$  of  $G$  such that the moment map of  $\psi'$  with respect to the representation  $\Phi'$  is equal to the *negation* of  $\omega$ , i.e.,

$$\Omega_{\Phi'}(\psi') = -\omega. \quad (4.3.1)$$

Since moment maps are additive across tensor products ([Lemma 3.2.39](#)), we can conclude that the moment map of the unknown state  $\psi$  equals  $\omega$  if and only if the moment map for the tensor product of  $\psi$  and  $\psi'$  vanishes,

$$\Omega_\Phi(\psi) = \omega \iff \Omega_{\Phi \otimes \Phi'}(\psi \otimes \psi') = \Omega_\Phi(\psi) + \Omega_{\Phi'}(\psi') = 0. \quad (4.3.2)$$

Consequently, it becomes possible to apply [Corollary 3.3.21](#) to the tensor product representation and obtain an asymptotic test for whether or not  $\Omega_\Phi(\psi) = \omega$ . These observations bring us to the following extension of [Theorem 3.3.20](#), which we refer to henceforth as the *inversion trick*.

**Proposition 4.3.1.** *Let  $\Phi : G \rightarrow \text{GL}(\mathcal{V})$  and  $\Phi' : G \rightarrow \text{GL}(\mathcal{V}')$  be representations of a complex reductive group  $G$ . Furthermore, let  $v' \in V'$  be a non-zero vector such that*

$$\Omega_{\Phi'}([v']) = -\omega \in i\mathfrak{k}^*. \quad (4.3.3)$$

Let  $E_\omega^n$  be defined by

$$E_\omega^n := \text{Tr}_{\mathcal{V}^{\otimes n}} \left( \Gamma_{(\Phi \otimes \Phi')^{\otimes n}} P_{[v']}^{\otimes n} \right). \quad (4.3.4)$$

where  $\Gamma_{(\Phi \otimes \Phi')^{\otimes n}}$  is the projection operator onto the subspace of vectors in  $(\mathcal{V} \otimes \mathcal{V}')^{\otimes n}$  invariant under the action of  $G$ . Then for all  $\psi \in \mathbb{P}\mathcal{H}$ ,

$$\begin{aligned} \Omega_\Phi(\psi) = \omega &\implies \limsup_{n \rightarrow \infty} \text{Tr}(P_\psi^{\otimes n} E_\omega^n)^{\frac{1}{n}} = 1, \\ \Omega_\Phi(\psi) \neq \omega &\implies \limsup_{n \rightarrow \infty} \text{Tr}(P_\psi^{\otimes n} E_\omega^n)^{\frac{1}{n}} < 1. \end{aligned} \quad (4.3.5)$$

*Proof.* Let  $v \in \mathcal{V}$  and  $v' \in \mathcal{V}'$  be unit vectors such that  $[v] = \psi \in \mathbb{P}\mathcal{V}$  and  $[v'] = \psi' \in \mathbb{P}\mathcal{H}$ . An application of [Theorem 3.3.20](#) to  $\Phi \otimes \Phi' : G \rightarrow \mathrm{GL}(\mathcal{V} \otimes \mathcal{W})$  yields

$$\limsup_{n \rightarrow \infty} \left\| \Gamma_{(\Phi \otimes \Phi')^{\otimes n}}(v^{\otimes n} \otimes v'^{\otimes n}) \right\|^{\frac{1}{n}} = \mathrm{cap}_{\Phi \otimes \Phi'}(v \otimes v') = \inf_{g \in G} \|\Phi(g)v\| \|\Phi'(g)v'\|, \quad (4.3.6)$$

and therefore, because  $v$  and  $v'$  are assumed unit vectors,

$$\limsup_{n \rightarrow \infty} \mathrm{Tr}(P_\psi^{\otimes n} E_{\nu'}^n)^{\frac{1}{n}} = \mathrm{cap}_{\Phi \otimes \Phi'}^2(v \otimes v'). \quad (4.3.7)$$

By the Kempf-Ness theorem [Theorem 3.2.31](#), the capacity  $\mathrm{cap}_{\Phi \otimes \Phi'}(v \otimes v')$  is maximized and equal to one if and only if  $\Omega_{\Phi \otimes \Phi'}(\psi \otimes \psi') = 0$  or equivalently,

$$\Omega_\Phi(\psi) = -\Omega_{\Phi'}(\psi') = \nu, \quad (4.3.8)$$

which proves the claim.  $\square$

### 4.3.2 Deforming moment maps

The purpose this section is to introduce a more sophisticated variation of the inversion trick from [Section 4.3.1](#), which we call the *deformation trick*. The deformation trick has both a geometric part, which modifies the moment map by an affine translation, a coadjoint evolution, and a rescaling by a positive integer, and an invariant part relating invariant subspaces of one representation to the fixed subspaces of another. This trick is also sometimes called the *shifting trick* [[Bri87](#); [MN84](#)].

In any case, in order to apply these tricks, it will be useful to have a source of reference vectors whose moment maps are well understood.

**Lemma 4.3.2.** *Let  $\Phi : G \rightarrow \mathrm{GL}(\mathcal{V})$  be a representation of a complex reductive group  $G$  with Lie algebra  $\mathfrak{g} = \mathfrak{k}_\mathbb{C}$  and let  $v_\lambda \in \mathcal{V}$  be a weight vector for  $\Phi$  with weight  $\lambda \in i\mathfrak{k}^*$ <sup>3</sup>. Then the moment map evaluated on the ray  $[v_\lambda] \in \mathbb{P}\mathcal{V}$  generated by  $v_\lambda$  can be identified with the weight  $\lambda$ ,*

$$\Omega_\Phi([v_\lambda]) = \lambda. \quad (4.3.9)$$

*Proof.* Recall the root-space decomposition of a complex semisimple Lie algebra  $\mathfrak{g} = \mathfrak{k}_\mathbb{C}$  takes the form of an orthogonal, direct-sum decomposition,

$$\mathfrak{g} = \mathfrak{h} \oplus \bigoplus_{\alpha \in R} \mathfrak{g}_\alpha. \quad (4.3.10)$$

where  $R$  is the root system of roots and  $\mathfrak{g}_\alpha$  is the root space with root  $\alpha \in \mathfrak{h}^*$  and  $\mathfrak{h} = \mathfrak{g}_0 = \mathfrak{k}_\mathbb{C}$  is the Cartan subalgebra. If  $v_\lambda \in \mathcal{V}_\times$  is a weight vector with respect to  $\phi : \mathfrak{g} \rightarrow \mathfrak{gl}(\mathcal{V})$  of weight  $\lambda \in \mathfrak{h}^*$ , then  $\phi(H)v_\lambda = \lambda(H)v_\lambda$  for all  $H \in \mathfrak{h}$ . Therefore,  $\Omega([v_\lambda])(H) = \lambda(H)$  for all  $H \in \mathfrak{h}$  which covers all  $H \in i\mathfrak{t} \subset \mathfrak{h}$ . Now consider  $\Omega([v_\lambda])(X)$  for  $X \notin \mathfrak{h}$ , where

$$\Omega([v_\lambda])(X) = \frac{\langle v_\lambda, \phi(X)v_\lambda \rangle}{\langle v_\lambda, v_\lambda \rangle}. \quad (4.3.11)$$

---

<sup>3</sup>Viewed as an element of the subspace  $i\mathfrak{t}^* \subset i\mathfrak{k}^*$  in the sense of [Remark 2.4.50](#).



If  $X \in \mathfrak{g}_\alpha$  is a root vector with non-zero root  $\alpha \in R$ , then  $\phi(X)v_\lambda$  is either zero or a weight vector with weight  $\lambda + \alpha$  because for all  $H \in \mathfrak{h}$ ,

$$\phi(H)\phi(X)v_\lambda = \phi([H, X])v_\lambda + \phi(X)\phi(H)v_\lambda = (\alpha(H) + \lambda(H))\phi(X)v_\lambda. \quad (4.3.12)$$

If  $\phi(X)v_\lambda$  is zero, then  $\Omega([v_\lambda])(X)$  is obviously zero. If  $\phi(X)v_\lambda$  is non-zero, then it is a weight vector of weight  $\alpha + \lambda$  yet  $\Omega([v_\lambda])(X)$  is also zero because weight spaces of distinct weights are orthogonal to each other and  $\phi(X)v_\lambda \in \mathcal{V}_{\alpha+\lambda}$ , while  $v_\lambda \in \mathcal{V}_\lambda$ . In summary, the moment map  $\Omega([v_\lambda]) \in (\mathfrak{it}^*)^*$  satisfies  $\Omega([v_\lambda])(H) = \lambda(H)$  for  $H \in \mathfrak{it}$  and  $\Omega([v_\lambda])(X) = 0$  for  $X$  orthogonal to  $\mathfrak{it}$ . Therefore,  $\Omega([v_\lambda])$  can be identified with the weight  $\lambda \in \mathfrak{it} \subseteq \mathfrak{it}^*$  as claimed above.  $\square$

Unfortunately, while [Lemma 4.3.2](#) provides reference vectors  $v$  which have known moment maps, it does not exhaust all of the possible elements of  $\mathfrak{it}^*$  that one might wish to find a reference vector for in the context of the [Proposition 4.3.1](#). This is because there are only a finite number of weights in any given finite dimensional representation, and thus the set of all weights of a representation does not exhaust all of the possible elements in  $\mathfrak{it}^*$ . Fortunately, using the various symmetries of moment maps, i.e. [Lemma 3.2.39](#) and [Lemma 3.2.35](#), it becomes possible to relate every point  $\nu \in \mathfrak{it}^*$  which is a rational multiple of some element in the coadjoint orbit of a dominant analytically integral element to the moment map of a fixed vector in a natural way.

**Definition 4.3.3.** Let  $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{it}^*$  be the Lie algebra of a complex reductive group  $G = K_{\mathbb{C}}$  with fixed maximal abelian subalgebra  $\mathfrak{h} = \mathfrak{t} \oplus \mathfrak{it}$ . Further suppose that a base for the root system  $R \subseteq \mathfrak{it}^* \subseteq \mathfrak{it}^*$  of  $\mathfrak{g}$  relative to  $\mathfrak{h}$  is chosen and let  $\mathfrak{it}_+^* \subseteq \mathfrak{it}^*$  be the positive Weyl chamber. Also let  $\Lambda \subseteq \mathfrak{it}^*$  be the set of analytically integral elements (with respect to the torus  $T \subseteq K$ ) and let  $\Lambda_+ = \Lambda \cap \mathfrak{it}_+^*$  be the set of dominant, analytically integral elements. It is a well-known result that the coadjoint orbit of a given  $\omega \in \mathfrak{it}^*$  intersects  $\mathfrak{it}_+^*$  uniquely. Let  $\omega_+ \in \mathfrak{it}_+^*$  be this unique intersection point and let  $h \in K$  be such that

$$\omega = \text{Ad}^*(h)(\omega_+) = h \cdot \omega_+. \quad (4.3.13)$$

An element  $\omega \in \mathfrak{it}^*$  is said to have a **rational** coadjoint orbit if there exists a positive integer  $\ell \in \mathbb{N}$ , a group element  $h \in K$ , and a dominant, analytically integral element  $\lambda \in \Lambda_+$  such that

$$\ell\omega = \text{Ad}^*(h)(\lambda) = h \cdot \lambda. \quad (4.3.14)$$

If  $\omega \in \mathfrak{it}^*$  has a rational coadjoint orbit, the value of  $\lambda$  which satisfies the above equation is uniquely determined by taking  $\ell$  to be as small as possible (in which case  $\lambda = (\ell\omega)_+$ ).

The main reason for considering  $\omega \in \mathfrak{it}^*$  with rational coadjoint orbits is the following result which generalizes the key ingredient of the deformation trick.

**Lemma 4.3.4.** *Let everything be as in [Definition 4.3.3](#). Let  $\omega \in \mathfrak{it}^*$  have a rational coadjoint orbit such that  $\ell\omega = \text{Ad}^*(h)(\lambda)$  holds for some positive integer  $\ell \in \mathbb{N}$ ,  $h \in K$  and dominant, analytically integral element  $\lambda \in \Lambda_+ \subset \mathfrak{it}^*$ . Let  $\Phi_\lambda : G \rightarrow \text{GL}(\mathcal{V}_\lambda)$  be the irreducible representation of  $G = K_{\mathbb{C}}$  with highest weight  $\lambda$  and let  $v_\lambda \in \mathcal{V}_\lambda$  be a highest weight vector.*

Furthermore, let  $\Phi : G \rightarrow \mathrm{GL}(\mathcal{V})$  be a representation on  $\mathcal{V}$  and let  $v \in \mathcal{V}$  be a vector. Then the moment map of  $[v] \in \mathbb{P}\mathcal{V}$  equals

$$\Omega_{\Phi}([v]) = \omega = \ell^{-1} \mathrm{Ad}^*(h)(\lambda), \quad (4.3.15)$$

if and only if the moment map of  $v^{\otimes \ell} \otimes \Phi_{\lambda}^*(h)v_{\lambda}^*$  with respect to the representation  $\Phi^{\otimes \ell} \otimes \Phi_{\lambda}^*$  vanishes, i.e.,

$$\Omega_{\Phi^{\otimes \ell} \otimes \Phi_{\lambda}^*}([v^{\otimes \ell} \otimes \Phi_{\lambda}^*(h)v_{\lambda}^*]) = 0. \quad (4.3.16)$$

*Proof.* Since  $v_{\lambda}$  is a weight vector of weight  $\lambda$  with respect to  $\Phi_{\lambda}$ , its dual  $v_{\lambda}^* := \langle v_{\lambda}, \cdot \rangle \in \mathcal{V}_{\lambda}^*$  is a weight vector with respect to the dual representation  $\Phi_{\lambda}^*$  with weight  $-\lambda$ . Then, using [Lemma 3.2.39](#) and [Lemma 4.3.2](#), the moment map of the representation

$$\Phi^{\otimes \ell} \otimes \Phi_{\lambda}^* : G \rightarrow \mathrm{GL}(\mathcal{V}^{\otimes \ell} \otimes \mathcal{V}_{\lambda}^*) \quad (4.3.17)$$

is related to the moment map of  $\Phi : G \rightarrow \mathrm{GL}(\mathcal{V})$  via

$$\Omega_{\Phi^{\otimes \ell} \otimes \Phi_{\lambda}^*}([v^{\otimes \ell} \otimes v_{\lambda}^*]) = \ell \Omega_{\Phi}([v]) + \Omega_{\Phi_{\lambda}^*}([v_{\lambda}^*]) = \ell \Omega_{\Phi}([v]) - \lambda. \quad (4.3.18)$$

Furthermore, by [Lemma 3.2.35](#), we have

$$\frac{1}{\ell} \Omega_{\Phi^{\otimes \ell} \otimes \Phi_{\lambda}^*}([v^{\otimes \ell} \otimes \Phi_{\lambda}^*(h)v_{\lambda}^*]) = \Omega_{\Phi}([v]) - \ell^{-1} \mathrm{Ad}^*(h)\lambda = \Omega_{\Phi}([v]) - \omega, \quad (4.3.19)$$

which proves the claim.  $\square$

**Remark 4.3.5.** At this stage it is worth clarifying that if  $v_{\lambda} \in \mathcal{V}_{\lambda}$  is a highest weight vector of the highest weight representation  $\Phi : G \rightarrow \mathrm{GL}(\mathcal{V}_{\lambda})$ , then  $v_{\lambda}^* := \langle v_{\lambda}, \cdot \rangle \in \mathcal{V}_{\lambda}^*$  is a weight vector of the dual representation  $\Phi_{\lambda}^* : G \rightarrow \mathrm{GL}(\mathcal{V}_{\lambda}^*)$ , albeit with opposite weight  $-\lambda$ , as

$$\Phi_{\lambda}^*(g)v_{\lambda}^* = \Phi_{\lambda}^*(g)(\langle v_{\lambda}, \cdot \rangle) = \langle v_{\lambda}, \Phi_{\lambda}(g^{-1}) \cdot \rangle = \langle \Phi_{\lambda}(g^{-1})^* v_{\lambda}, \cdot \rangle \in \mathcal{V}_{\lambda}^*. \quad (4.3.20)$$

Moreover,  $v_{\lambda}^*$  is *not* the highest weight vector in  $\mathcal{V}_{\lambda}^*$ , but instead the *lowest* weight vector in  $\mathcal{V}_{\lambda}^*$ . To obtain the *highest* weight of the dual representation, one needs to define an involution on the weights by sending  $\lambda$  to  $\lambda^* := -w_0(\lambda)$  where  $w_0$  is the unique longest element in the Weyl group. Then  $\lambda^*$  will be the highest weight in the dual representation,  $\mathcal{V}_{\lambda}^*$ , and therefore representation  $\mathcal{V}_{\lambda^*}$  with highest weight  $\lambda^*$  (dual to  $\lambda$ ) is isomorphic to the dual representation, such that  $\mathcal{V}_{\lambda^*} \cong \mathcal{V}_{\lambda}^*$ . Both of these options have been used to define a deformation of the moment map, e.g, by  $\lambda^*$  in [\[Bür+19\]](#) and by  $-\lambda$  in [\[FW20\]](#).

### 4.3.3 Rotating & scaling extremal weight vectors

The aim of this section will be to describe how the action of the complexification  $G = K_{\mathbb{C}}$  of a compact connected Lie group  $K$  on a highest weight vector  $v_{\lambda} \in \mathcal{H}_{\lambda}$  can always be decomposed into a rotation of  $v_{\lambda}$  through its  $K$ -orbit and then a scaling of norm by a scalar factor which depends on the weight  $\lambda$ .

**Theorem 4.3.6.** *Let  $K$  be a compact, connected Lie group with  $G = K_{\mathbb{C}}$  its complexification and  $T \subseteq K$  be a fixed maximal torus with Lie algebra  $\mathfrak{t}$ . Let  $\Phi_{\lambda} : G \rightarrow \mathrm{GL}(\mathcal{H}_{\lambda})$  be an irreducible representation of  $G$  with highest weight  $\lambda : \mathfrak{it} \rightarrow \mathbb{R}$  and let  $v_{\lambda} \in \mathcal{H}_{\lambda}$  be a highest weight vector. Then there exists maps  $\alpha_{+} : G \rightarrow \mathfrak{it}$  and  $\kappa_{+} : G \rightarrow K$  such that the action of  $G$  on  $v_{\lambda}$  satisfies*

$$\Phi_{\lambda}(g)v_{\lambda} = e^{\lambda(\alpha_{+}(g))}\Phi_{\lambda}(\kappa_{+}(g))v_{\lambda}. \quad (4.3.21)$$

*In other words,  $g$  acts on  $K$ -orbits of highest weight vectors by scaling by the factor  $e^{\lambda(\alpha_{+}(g))} \in \mathbb{R}$ , since*

$$\|\Phi_{\lambda}(g)v_{\lambda}\| = e^{\lambda(\alpha_{+}(g))} \|v_{\lambda}\|. \quad (4.3.22)$$

*Proof.* The proof relies on the Iwasawa decomposition of  $G$  previously covered in [Proposition 2.4.60](#). Let  $B$  be a fixed Borel subgroup containing the maximal torus  $T$ , and let the unique Iwasawa decomposition of  $g \in G$  be  $g = k_g \cdot a_g \cdot n_g$ . Since  $a_g \in A = \exp(\mathfrak{it})$  is uniquely determined by  $g$ , let  $\alpha_{+} : G \rightarrow \mathfrak{a}$  be the map satisfying  $\exp(\alpha_{+}(g)) = a_g$ . Additionally, let  $\kappa_{+} : G \rightarrow K$  be the map satisfying  $\kappa_{+}(g) = k_g$ . Now, by the definition of the highest weight vector  $v_{\lambda}$ , the action of positive root vector  $X \in \mathfrak{g}_{\nu}$  with positive root  $\nu$  satisfies  $\phi_{\lambda}(X)v_{\lambda} = 0$ . Therefore, the nilpotent subgroup  $N = [B, B]$  with Lie algebra  $\mathfrak{n}$  generated by all positive roots, satisfies for all  $n \in N$ ,

$$\Phi_{\lambda}(n)v_{\lambda} = v_{\lambda}. \quad (4.3.23)$$

Furthermore, as  $\alpha_{+}(g) \in \mathfrak{a} = \mathfrak{it} \subseteq \mathfrak{t} \oplus \mathfrak{it} = \mathfrak{h}$ , and  $a_g = \exp(\alpha(g))$ , the highest weight vector  $v_{\lambda}$  is an eigenvector of  $\Phi_{\lambda}(a_g)$  with eigenvalue  $e^{\lambda(\alpha_{+}(g))} \in \mathbb{R}$ , i.e.

$$\Phi_{\lambda}(a_g)v_{\lambda} = \Phi_{\lambda}(\exp(\alpha_{+}(g)))v_{\lambda} = e^{\phi_{\lambda}(\alpha_{+}(g))}v_{\lambda} = e^{\lambda(\alpha_{+}(g))}v_{\lambda}. \quad (4.3.24)$$

Therefore,

$$\Phi_{\lambda}(g)v_{\lambda} = \Phi_{\lambda}(k_g)\Phi_{\lambda}(a_g)\Phi_{\lambda}(n_g)v_{\lambda} = e^{\lambda(\alpha_{+}(g))}\Phi_{\lambda}(\kappa_{+}(g))v_{\lambda}. \quad (4.3.25)$$

□

**Corollary 4.3.7.** *Let everything be defined as in [Theorem 4.3.6](#). Then there exists maps  $\alpha_{-} : G \rightarrow \mathfrak{it}$  and  $\kappa_{-} : G \rightarrow K$  such that*

$$\Phi_{\lambda}^{*}(g)v_{\lambda}^{*} = e^{\lambda(\alpha_{-}(g))}\Phi_{\lambda}^{*}(\kappa_{-}(g))v_{\lambda}^{*}, \quad (4.3.26)$$

*and therefore*

$$\|\Phi_{\lambda}^{*}(g)v_{\lambda}^{*}\| = e^{\lambda(\alpha_{-}(g))} \|v_{\lambda}^{*}\|. \quad (4.3.27)$$

*Proof.* The proof proceeds in exactly the same way to the proof of [Theorem 4.3.6](#). The only difference is to take the Iwasawa decomposition of  $g \in G$  with respect to the so-called opposite Borel subgroup  $B_{-} \subset G$  with maximal unipotent subgroup  $N_{-}$  with Lie algebra generated by all negative roots. The reason for this difference lies with the fact that  $v_{\lambda}^{*}$  is a lowest weight vector for the dual representation (see [Remark 4.3.5](#)). □

While the scalar factor above depends on the highest weight  $\lambda \in \mathfrak{it}^{*}$ , it can be extended to a function which is well-defined for all elements  $\omega \in \mathfrak{it}^{*}$ .

**Definition 4.3.8.** Let  $G = K_{\mathbb{C}}$  be a complex reductive group with maximal compact subgroup  $K$  with fixed Borel subgroup  $B$  and maximal torus  $T = B \cap K$  with Lie algebra  $\mathfrak{t}$  and let  $i\mathfrak{t}_+^*$  be the closure of the fundamental Weyl chamber  $i\mathfrak{t}_+^* \subseteq \mathfrak{t}^*$ . The coadjoint orbit of  $\omega \in i\mathfrak{k}^*$ , denoted by  $K \cdot \omega = \text{Ad}^*(K)(\omega) \subseteq i\mathfrak{k}^*$ , intersects  $i\mathfrak{t}_+^*$  at a unique point  $\omega_+ \in i\mathfrak{t}_+^*$ . Let  $h \in K$  be such that  $\omega = \text{Ad}^*(h)(\omega_+)$ . Furthermore, let  $\alpha_- : G \rightarrow i\mathfrak{t}$  be the map from [Corollary 4.3.7](#). Then define the function,

$$\chi_\omega : G \rightarrow (0, \infty), \quad (4.3.28)$$

for all  $g \in G$  by

$$\chi_\omega(g) := e^{\omega_+(\alpha_-(gh))}. \quad (4.3.29)$$

**Remark 4.3.9.** Note that the function  $\chi_\omega : G \rightarrow (0, \infty)$  has been explicitly defined such that if  $\omega \in i\mathfrak{k}^*$  has a rational coadjoint orbit, meaning there exists an  $\ell \in \mathbb{N}$  and dominant, analytically integral element  $\lambda$  such that  $\omega = \ell^{-1}\text{Ad}^*(h)(\lambda)$  (and thus  $\ell\omega_+ = \lambda$ ), then  $\chi_\omega(g)$  can be expressed as

$$\chi_\omega(g) = \|\Phi_\lambda^*(gh)v_\lambda\|^{\frac{1}{\ell}} = e^{\ell^{-1}\lambda(\alpha_-(gh))}, \quad (4.3.30)$$

where  $\Phi_\lambda : G \rightarrow \text{GL}(\mathcal{H}_\lambda)$  is an irreducible representation with highest weight  $\lambda$  with highest weight vector unit vector  $v_\lambda$  ( $\|v_\lambda\| = 1$ ).

### 4.3.4 Deformed capacity

In [Section 4.3.2](#), we saw how to transform the moment map of one representation to the moment map of another representation. In this section, we consider happens to the capacity of a vector with respect to the same transformation of representations.

**Definition 4.3.10.** Let  $\Phi : G \rightarrow \text{GL}(\mathcal{H})$  be a representation of  $G$  and let  $v \in H$ . Let  $\omega \in i\mathfrak{k}^*$  be arbitrary. Define the  **$\omega$ -capacity** of  $v$  as

$$\text{cap}_\Phi^\omega(v) := \inf_{g \in G} \chi_\omega(g) \|\Phi(g)v\|. \quad (4.3.31)$$

Alternatively, we say  $\text{cap}_\Phi^\omega(v)$  is the capacity of  $v$  *deformed* by  $\omega$ .

**Proposition 4.3.11.** *The  $\omega$ -capacity of a vector  $v$  as defined by [Definition 4.3.10](#) satisfies*

$$\text{cap}_\Phi^\omega(v) = \|v\| \iff \Omega_\Phi([v]) = \omega. \quad (4.3.32)$$

*If  $\omega$  has a rational coadjoint orbit, i.e.  $\ell\omega = \text{Ad}^*(h)(\lambda)$ , then by [Remark 4.3.9](#), we have*

$$\begin{aligned} \text{cap}_\Phi^\omega(v) &= \text{cap}_{\Phi^{\otimes \ell} \otimes \Phi_\lambda^*}(v^{\otimes \ell} \otimes \Phi_\lambda^*(h)v_\lambda^*)^{\frac{1}{\ell}}, \\ &= \text{cap}_{\Phi^{\otimes \ell} \otimes \Phi_\lambda^*}((\Phi(h^{-1})v)^{\otimes \ell} \otimes v_\lambda^*)^{\frac{1}{\ell}}. \end{aligned} \quad (4.3.33)$$

*Proof.* The claim follows from [Lemma 4.3.4](#), the density of rational coadjoint orbits in  $i\mathfrak{k}^*$  ([Definition 4.3.3](#)) and the continuity of the moment map.  $\square$

**Proposition 4.3.12.** *Let  $\Phi : G \rightarrow \mathrm{GL}(\mathcal{H})$  be a representation of  $G = K_{\mathbb{C}}$  on  $\mathcal{H}$ . Let  $\lambda \in \Lambda_+$  be a dominant, analytically integral weight, let  $n \in \mathbb{N}$  be a positive integer and let  $h \in K$ . Define  $\omega = n^{-1}\mathrm{Ad}^*(h)(\lambda)$ . Then for all  $v \in \mathcal{H}$ ,*

$$\left\| \Pi_{\Phi^{\otimes n}}^{\lambda}(\Phi(h^{-1})v)^{\otimes n} \right\|^{\frac{1}{n}} \leq \mathrm{cap}_{\Phi}^{\omega}(v), \quad (4.3.34)$$

where  $\Pi_{\Phi^{\otimes n}}^{\lambda}$  is the projection operator onto the subspace of highest weight vectors in  $\mathcal{H}^{\otimes n}$ .

*Proof.* The proof relies on the same technique from [Lemma 3.3.19](#) along with [Corollary 4.3.7](#). For all  $g \in G$ ,

$$\left\| \Pi_{\Phi^{\otimes n}}^{\lambda}(\Phi(h^{-1})v)^{\otimes n} \right\| = \left\| \Pi_{\Phi^{\otimes n}}^{\lambda} \Phi((gh)^{-1})^{\otimes n} \Phi(g)^{\otimes n} v^{\otimes n} \right\|, \quad (4.3.35)$$

$$= e^{\lambda(\alpha_-(gh))} \left\| \Pi_{\Phi^{\otimes n}}^{\lambda} \Phi(\kappa_-(g)^{-1}) \Phi(g)^{\otimes n} v^{\otimes n} \right\|, \quad (4.3.36)$$

$$= e^{\lambda(\alpha_-(gh))} \left\| \Phi(\kappa_-(g)) \Pi_{\Phi^{\otimes n}}^{\lambda} \Phi(\kappa_-(g)^{-1}) \Phi(g)^{\otimes n} v^{\otimes n} \right\|, \quad (4.3.37)$$

$$\leq e^{\lambda(\alpha_-(gh))} \left\| \Phi(g)^{\otimes n} v^{\otimes n} \right\|, \quad (4.3.38)$$

$$= \left( e^{n^{-1}\lambda(\alpha_-(gh))} \left\| \Phi(g)v \right\| \right)^n. \quad (4.3.39)$$

Since  $\chi_{\omega}(g) = e^{n^{-1}\lambda(\alpha_-(gh))}$  when  $\omega = n^{-1}\mathrm{Ad}^*(h)(\lambda)$ , it has been proven that

$$\left\| \Pi_{\Phi^{\otimes n}}^{\lambda}(\Phi(h^{-1})v)^{\otimes n} \right\|^{\frac{1}{n}} = (\chi_{\omega}(g) \left\| \Phi(g)v \right\|)^n. \quad (4.3.40)$$

Optimizing over all  $g \in G$  yields the result.  $\square$

The lower bound associated to [Proposition 4.3.12](#), proven next, is slightly harder to state as it relies on replacing the  $\lambda$  appearing in [Proposition 4.3.12](#) with a dominant weight that grows proportionally with increasing  $n$ . For this we require the following definition.

**Definition 4.3.13.** Let  $\omega \in i\mathfrak{k}^*$  have rational coadjoint orbit, meaning there exists a dominant, analytically integral element  $\lambda \in \Lambda_+ \in i\mathfrak{k}^*$ , an element  $h \in K$ , and a positive integer  $\ell \in \mathbb{N}$  (which may be assumed as small as possible) such that

$$\ell\omega = \mathrm{Ad}^*(h)(\lambda), \quad (4.3.41)$$

and  $\omega_+ := \lambda/\ell \in \mathfrak{k}^*$  is dominant. Given a representation  $\Phi : G \rightarrow \mathrm{GL}(\mathcal{H})$  of  $G = K_{\mathbb{C}}$  and a positive integer  $n \in \mathbb{N}$ , define  $\Pi_{\Phi^{\otimes n}}^{n\omega_+, h}$  to be the projection operator onto the subspace of highest weight vectors in  $\mathcal{H}^{\otimes n}$  with weight  $n\omega_+ = n\lambda/\ell$ , albeit rotated by  $h \in K$ :

$$\Pi_{\Phi^{\otimes n}}^{n\omega_+, h} := \Phi^{\otimes n}(h) \Pi_{\Phi^{\otimes n}}^{n\omega_+} \Phi^{\otimes n}(h^{-1}). \quad (4.3.42)$$

Note that if  $n\omega_+$  is not analytically integral for some value of  $n$ , then  $\Pi_{\Phi^{\otimes n}}^{n\omega_+} = 0$  and thus  $\Pi_{\Phi^{\otimes n}}^{n\omega_+, h} = 0$ .

**Theorem 4.3.14.** *Let everything be defined as in [Definition 4.3.13](#). Then for all vectors  $v \in \mathcal{V}$ ,*

$$\limsup_{n \rightarrow \infty} \left\| \Pi_{\Phi^{\otimes n}}^{n\omega_+, h} v^{\otimes n} \right\|^{\frac{1}{n}} = \mathrm{cap}_{\Phi}^{\omega}(v). \quad (4.3.43)$$

*Proof.* The proof relies on the original strong duality result [Theorem 3.3.20](#) from [Section 3.3.4](#). From [Theorem 3.3.20](#) we get

$$\text{cap}_{\Phi^{\otimes \ell} \otimes \Phi_\lambda^*}(v^{\otimes \ell} \otimes \Phi_\lambda^*(h)v_\lambda^*) = \limsup_{k \rightarrow \infty} \left\| \Gamma_{(\Phi^{\otimes \ell} \otimes \Phi_\lambda^*)^{\otimes k}}((\Phi(h^{-1})v)^{\otimes \ell} \otimes v_\lambda^*)^{\otimes k} \right\|^{\frac{1}{k}}, \quad (4.3.44)$$

where  $\Gamma_{(\Phi^{\otimes \ell} \otimes \Phi_\lambda^*)^{\otimes k}}$  is the projection operator onto the subspace of  $G$ -invariant vectors in  $(\mathcal{H}^{\otimes \ell} \otimes \mathcal{V}_\lambda^*)^{\otimes k}$ . By [Proposition 4.3.11](#), we can relate the  $\omega$ -capacity of  $v$  to the capacity above to obtain

$$\text{cap}_\Phi^\omega(v) = \limsup_{k \rightarrow \infty} \left\| \Gamma_{(\Phi^{\otimes \ell} \otimes \Phi_\lambda^*)^{\otimes k}}((\Phi(h^{-1})v)^{\otimes \ell} \otimes v_\lambda^*)^{\otimes k} \right\|^{\frac{1}{k\ell}}, \quad (4.3.45)$$

Since  $v_\lambda$  is a weight-vector of highest weight  $\lambda$  in  $\mathcal{V}_\lambda$ ,  $v_\lambda^{\otimes k}$  is a vector of highest weight  $k\lambda$  in  $\mathcal{V}_{k\lambda} \subseteq \mathcal{V}_\lambda^{\otimes k}$  and thus Schur's lemma implies

$$\left\| \Gamma_{(\Phi^{\otimes \ell} \otimes \Phi_\lambda^*)^{\otimes k}}((\Phi(h^{-1})v)^{\otimes \ell} \otimes v_\lambda^*)^{\otimes k} \right\| = \frac{1}{\sqrt{\dim(\mathcal{V}_{k\lambda})}} \left\| \Pi_{\Phi^{\otimes k\ell}}^{k\lambda}(\Phi(h^{-1})v)^{\otimes k\ell} \right\|. \quad (4.3.46)$$

Let  $n = k\ell$  be such that  $k\lambda = k\ell\omega_+ = n\omega_+$ . Then, taking the appropriate limit as  $n \rightarrow \infty$  yields the result because  $\dim(\mathcal{V}_{k\lambda})$  only grows polynomially with increasing  $k$  and thus does not affect the value of the limit.  $\square$

Consider the asymptotics of the sequence of probabilities,  $\left\| \Pi_{\Phi^{\otimes n}}^\lambda(\Phi(h^{-1})v)^{\otimes n} \right\|^2$ , in the statement of [Theorem 3.3.20](#) (where  $v$  is assumed a unit vector). From the statement of [Theorem 3.3.20](#), we observe that this sequence of probabilities decays to zero at an exponential rate with increasing  $n$  given by  $I_v(\mu) := -\log \text{cap}_\Phi^\mu(v)^2$ . Moreover, by [Proposition 4.3.11](#), this rate vanishes if and only if  $\Omega_\Phi([v]) = \mu$ . In [Section 4.4.1](#), we see how to bundle together the projection operators  $\Phi(h)^{\otimes n} \Pi_{\Phi^{\otimes n}}^\lambda \Phi(h^{-1})^{\otimes n}$  in [Theorem 3.3.20](#) to obtain a positive operator valued measure which concentrates around  $\Omega_\Phi([v])$  for each unit vector  $v \in \mathcal{H}$ .

### 4.3.5 Strong duality & purifications

An alternative formulation of [Theorem 3.3.20](#) is concerned with density operators  $\rho \in \mathcal{S}(\mathcal{H})$  rather than vectors  $v \in \mathcal{H}$ , and follows from recalling that every density operator admits of a purification living in a dilated vector space.

**Corollary 4.3.15.** *Let  $\Phi : G \rightarrow \text{GL}(\mathcal{H})$  be a representation of a complex reductive group  $G$ . Then for all density operators  $\rho \in \mathcal{S}(\mathcal{H})$ ,*

$$\inf_{g \in G} \text{Tr}(\rho \Phi(g^*g)) = \limsup_{n \rightarrow \infty} \text{Tr}(\rho^{\otimes n} \Gamma_{\Phi^{\otimes n}})^{\frac{1}{n}}, \quad (4.3.47)$$

where  $\Gamma_{\Phi^{\otimes n}}$  is the projective operator onto the subspace of  $\Phi^{\otimes n}$ -invariant vectors in  $\mathcal{H}^{\otimes n}$ .

*Proof.* The proof follows from an application of [Theorem 3.3.20](#) to a purification of the state  $\rho \in \mathcal{S}(\mathcal{H})$  alongside a lifted representation of  $G$  on the purifying space. Let  $\mathcal{Z} \cong \mathcal{H}$  be a vector space isomorphic to  $\mathcal{H}$  and let  $v \in \mathcal{H} \otimes \mathcal{Z}$  be a purification of  $\rho$ , i.e., a non-zero vector  $v$  such that for all operators  $L \in \text{End}(\mathcal{H})$

$$\text{Tr}_{\mathcal{H}}(\rho L) = \frac{\langle v, (L \otimes \mathbb{1}_{\mathcal{Z}})v \rangle}{\langle v, v \rangle} = \text{Tr}_{\mathcal{H} \otimes \mathcal{Z}}[(L \otimes \mathbb{1}_{\mathcal{Z}})P_{[v]}]. \quad (4.3.48)$$

Similarly, let  $\Phi' : G \rightarrow \text{GL}(\mathcal{H} \otimes \mathcal{Z})$  be the representation of  $G$  on  $\mathcal{V} \otimes \mathcal{Z}$  defined by

$$\Phi'(g) := \Phi(g) \otimes \mathbb{1}_{\mathcal{Z}}. \quad (4.3.49)$$

Then the capacity of the purifying vector  $v \in \mathcal{H} \otimes \mathcal{Z}$  with respect to  $\Phi'$  is

$$\text{cap}_{\Phi'}^2(v) = \inf_{g \in G} \|\Phi'(g)v\|^2, \quad (4.3.50)$$

$$= \inf_{g \in G} \langle \Phi'(g)v, \Phi'(g)v \rangle, \quad (4.3.51)$$

$$= \inf_{g \in G} \langle v, \Phi'(g^*g)v \rangle, \quad (4.3.52)$$

$$= \|v\|^2 \inf_{g \in G} \text{Tr}(\rho \Phi(g^*g)). \quad (4.3.53)$$

Moreover, if  $\Gamma_{\Phi^{\otimes n}} \in \text{End}(\mathcal{H}^{\otimes n})$  projects onto the subspace of  $G$ -invariant vectors with respect to  $\Phi$ , then  $\Gamma_{\Phi^{\otimes n}} \otimes \mathbb{1}_{\mathcal{Z}^{\otimes n}} \in \text{End}((\mathcal{H} \otimes \mathcal{Z})^{\otimes n})$  projects onto the subspace of  $G$ -invariant vectors with respect to  $\Phi'$ . Finally, as  $(\Gamma_{\Phi^{\otimes n}})^2 = \Gamma_{\Phi^{\otimes n}}$  is projective,

$$\|v\|^2 \text{Tr}(\rho^{\otimes n} \Gamma_{\Phi^{\otimes n}}) = \langle v^{\otimes n}, (\Gamma_{\Phi^{\otimes n}} \otimes \mathbb{1}_{\mathcal{Z}^{\otimes n}})v^{\otimes n} \rangle = \|(\Gamma_{\Phi^{\otimes n}} \otimes \mathbb{1}_{\mathcal{Z}^{\otimes n}})v^{\otimes n}\|^2. \quad (4.3.54)$$

An application of [Theorem 3.3.20](#) then implies the claimed result.  $\square$

In precisely the same way [Corollary 4.3.15](#) is a corollary from [Theorem 3.3.20](#), the following result is a corollary of [Theorem 4.3.14](#).

**Corollary 4.3.16.** *Let  $\Phi : G \rightarrow \text{GL}(\mathcal{H})$  be a representation of  $G$  on  $\mathcal{H}$  and suppose  $\omega \in i\mathfrak{k}^*$  has a rational coadjoint orbit. Then for all  $\rho \in \mathcal{S}(\mathcal{H})$ ,*

$$\inf_{g \in G} \chi_{\omega}^2(g) \text{Tr}(\Phi(g^*g)\rho) = \limsup_{n \rightarrow \infty} \text{Tr}(\Pi_{\Phi^{\otimes n}}^{n\omega_+, h} \rho^{\otimes n})^{\frac{1}{n}}, \quad (4.3.55)$$

where  $\chi_{\omega} : G \rightarrow (0, \infty)$  is given by [Definition 4.3.8](#), and  $\Pi_{\Phi^{\otimes n}}^{n\omega_+, h}$  is defined as in [Definition 4.3.13](#).

## 4.4 Estimation schemes

In [Section 4.4.1](#), we demonstrate how to estimate the moment map of a state  $\psi \in \mathbb{P}\mathcal{H}$  with respect to a given representation of a complex reductive group acting on  $\mathcal{H}$ . Afterwards, [Section 4.4.2](#) covers a more general framework of *property* estimation schemes.

### 4.4.1 Estimating moment maps

In this section, we will study the asymptotics of the measurement described by the POVM  $F_{\Phi^{\otimes n}}$  defined by [Eq. \(4.2.34\)](#) when the representation  $\Phi^{\otimes n} : G \rightarrow \text{GL}(\mathcal{H}^{\otimes n})$  is the  $n$ th tensor power representation of  $G$  on  $\mathcal{H}^{\otimes n}$ . As the highest weights of the representation  $\Phi^{\otimes n}$  scale proportionally with respect to  $n$ , it becomes more useful to define a *regularized* variant of the POVM  $F_{\Phi^{\otimes n}}$  by  $R_n^{\Phi}(\Delta) = F_{\Phi^{\otimes n}}(n\Delta)$ .



**Definition 4.4.1.** Let  $\Phi : G \rightarrow \text{GL}(\mathcal{H})$  be a representation of  $G = K_{\mathbb{C}}$  and let  $n \in \mathbb{N}$  be a positive integer. The *moment map estimation scheme* of order  $n$  is the POVM

$$R_n^\Phi : \Sigma(i\mathfrak{k}^*) \rightarrow \mathcal{B}(\mathcal{H}^{\otimes n}), \quad (4.4.1)$$

defined implicitly for all measurable functions  $g : i\mathfrak{k}^* \rightarrow \mathbb{R}$  and states  $\sigma_n \in \mathcal{S}(\mathcal{H}^{\otimes n})$  by

$$\int_{\omega \in i\mathfrak{k}^*} g(\omega) \text{Tr}(dR_n^\Phi(\omega)\sigma_n) = \sum_{\lambda \in \Lambda_+} d_\lambda \int_{k \in K} g(n^{-1}\text{Ad}^*(k)(\lambda)) \text{Tr}(\Pi_{\Phi^{\otimes n}}^{\lambda, k} \sigma_n) d\mu(k). \quad (4.4.2)$$

Whenever the POVM  $R_n^\Phi$  defined above is applied to  $n$  copies of a state  $\psi \in \mathbb{P}\mathcal{H}$ , the measurement outcomes in  $i\mathfrak{k}^*$  are distributed according to the probability measure,  $\xi_n^\psi : \Sigma(i\mathfrak{k}^*) \rightarrow [0, 1]$ , defined by

$$\xi_n^\psi(\Delta) = \text{Tr}(R_n^\Phi(\Delta)P_\psi^{\otimes n}). \quad (4.4.3)$$

The following result, [BCV21, Prop 3.24], proves that as  $n$  tends to infinity, the measurement outcomes,  $\omega \in i\mathfrak{k}^*$ , are concentrated around the value of the moment map for  $\psi$ ,  $\Omega_\Phi(\psi) \in i\mathfrak{k}^*$ .

**Theorem 4.4.2.** For each  $n \in \mathbb{N}$  let  $\xi_n^\psi : \Sigma(i\mathfrak{k}^*) \rightarrow [0, 1]$  be the probability measure in Eq. (4.4.3). Then the sequence  $(\xi_n^\psi)_{n \in \mathbb{N}}$  converges weakly to the Dirac measure  $\delta_{\Omega_\Phi(\psi)} : \Sigma(i\mathfrak{k}^*) \rightarrow [0, 1]$  concentrated on  $\Omega_\Phi(\psi) \in i\mathfrak{k}^*$ . Moreover, let  $I_\psi : i\mathfrak{k}^* \rightarrow [0, \infty]$  be the function defined for all  $\omega \in i\mathfrak{k}^*$  by

$$I_\psi(\omega) := -\log \text{cap}_\Phi^\omega(\psi)^2, \quad (4.4.4)$$

where  $\text{cap}_\Phi^\omega(v)$  is the  $\omega$ -capacity of  $v$  (Definition 4.3.10). Then the sequence of probability measures  $(\xi_n^\psi)_{n \in \mathbb{N}}$  satisfies the large deviation principle upper bound with rate function  $I_\psi$ . In fact, for all  $\Delta \in \Sigma(i\mathfrak{k}^*)$ ,

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log \xi_n^\psi(\Delta) \leq -\inf_{\omega \in \Delta} I_\psi(\omega). \quad (4.4.5)$$

*Proof.* From Proposition 4.3.12, we have, when  $\omega = n^{-1}\text{Ad}^*(k)(\lambda)$ , the inequality

$$\text{Tr}(\Phi(k)^{\otimes n} \Pi_{\Phi^{\otimes n}}^\lambda \Phi(k^{-1})^{\otimes n} P_\psi^{\otimes n})^2 = \|\Pi_{\Phi^{\otimes n}}^\lambda(\Phi(k^{-1})v)^{\otimes n}\|^2 \leq \text{cap}_\Phi^\omega(\psi)^{2n}. \quad (4.4.6)$$

Furthermore, the dimension  $d_\lambda = \dim(\mathcal{H}_\lambda)$  of any irreducible representation appearing in  $\mathcal{H}^{\otimes n}$  is at most the number of partitions of  $n$  into  $d = \dim(\mathcal{H})$  parts, we have the upper bound  $d_\lambda \leq (n+1)^{d(d-1)/2}$ . Therefore, since the Haar measure  $\mu$  on  $K$  is normalized,

$$\text{Tr}(R_n^\Phi(\Delta)P_\psi^{\otimes n}) \leq (n+1)^{d(d-1)/2} \sup_{\omega \in \Delta} \text{cap}_\Phi^\omega(\psi)^{2n}. \quad (4.4.7)$$

Taking the appropriate limit yields Eq. (4.4.5).

Moreover, note that  $I_\psi(\omega)$  vanishes if and only if  $\text{cap}_\Phi^\omega(v) = 1$  if and only if  $\Omega_\Phi(\psi) = \omega$  (Proposition 4.3.11). Therefore, Lemma 2.1.25 implies  $\xi_n^\psi$  converges weakly to the Dirac measure concentrated at the value of  $\Omega_\Phi(\psi)$ .  $\square$

**Remark 4.4.3.** The authors of Ref. [FW20] claim, without explicit proof, that in addition to the large deviation upper bound provided by Theorem 4.4.2, that the sequence of probability



measures  $(\xi_n^\psi)_{n \in \mathbb{N}}$  satisfies the large deviation principle *lower* bound with rate function  $I_\psi : i\mathfrak{k}^* \rightarrow [0, \infty]$ , i.e., for all open subsets  $O \in \Sigma(i\mathfrak{k}^*)$ ,

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \log \xi_n^\psi(O) \geq - \inf_{\omega \in O} I_\psi(\omega). \quad (4.4.8)$$

In light of [Theorem 4.3.14](#), such a result seems plausible. In Ref. [\[BCV21\]](#) however, the authors identify a number of additional assumptions which are sufficient to prove [Eq. \(4.4.8\)](#). For instance, if  $I_\psi : i\mathfrak{k}^* \rightarrow [0, \infty]$  happens to be a continuous function on its domain, defined as

$$\text{dom}(I_\psi) = \{\omega \in i\mathfrak{k}^* \mid I_\psi(\omega) < \infty\}, \quad (4.4.9)$$

then the proof technique from [\[BCV21\]](#) yields the lower bound of [Eq. \(4.4.8\)](#). The authors of Ref. [\[BCV21\]](#) additionally conjecture that this continuity condition for  $I_\psi$  indeed holds for all such rate functions, as it holds for all considered examples.

**Example 4.4.4.** Here we return to the example from [Section 4.1](#) for estimating the probability distribution  $q = (q_1, \dots, q_d)$  associated to the decomposition of a unit vector  $v \in \mathbb{C}^d$  into the standard orthonormal basis, i.e.,  $q_i = \|P_i v\|^2$  as in [Eq. \(4.1.7\)](#). In this context, we have  $K = \text{U}(1)^d$ ,  $G = \mathbb{C}_\times^d$  and  $i\mathfrak{k} = i\mathfrak{t} = \mathbb{R}^d$ , and the representation  $\Phi : \mathbb{C}_\times^d \rightarrow \text{GL}(d)$  defined by coordinate-wise multiplication as  $(\Phi(g_1, \dots, g_d)v)_k = g_k v_k$ . Then the function  $\alpha_- : \mathbb{C}_\times^d \rightarrow \mathbb{R}^d$  from [Corollary 4.3.7](#) satisfies

$$\alpha_-(g_1, \dots, g_d) = (\log |g_1|, \dots, \log |g_d|), \quad (4.4.10)$$

such that for any  $\omega \in (\mathbb{R}^d)^*$ , whose dual we identify with a  $p = (p_1, \dots, p_d) \in \mathbb{R}^d$ , the  $\omega$ -capacity of  $v$  satisfies

$$\text{cap}_\Phi^\omega(v)^2 = \inf_{g \in \mathbb{C}_\times^d} e^{-2\omega(\log |g_1|, \dots, \log |g_d|)} \sum_{j=1}^d |g_j v_j|^2, \quad (4.4.11)$$

$$= \inf_{x \in \mathbb{R}_{\geq 0}^d} e^{-2\omega(x_1, \dots, x_d)} \sum_{j=1}^d e^{2x_j} |v_j|^2, \quad (4.4.12)$$

$$= \inf_{x \in \mathbb{R}_{\geq 0}^d} e^{-2\langle p, x \rangle} \sum_{j=1}^d e^{2x_j} q. \quad (4.4.13)$$

Comparing this expression with [Eq. \(4.1.22\)](#), we recover  $I_q(p) = -\log \text{cap}_\Phi^\omega(v)^2$  as the relative entropy between  $p$  and  $q$  as expected.

**Example 4.4.5.** Here we consider the example of the quantum state estimation theorem due to Keyl [\[Key06, Thm. 3.2\]](#) (see also [\[BCV21, Ex. 3.17\]](#), and [\[FW20, Thm. 5.7\]](#)). Let  $K = \text{U}(d)$  be the group of unitary  $d \times d$  matrices such that  $i\mathfrak{k} = i\mathfrak{u}(d)$  is the set of all  $d \times d$  Hermitian matrices. Then  $G = K_\mathbb{C} = \text{GL}(d, \mathbb{C})$  is the group of invertible  $d \times d$  complex matrices. An element  $\omega = i\mathfrak{u}(d)^*$  then corresponds to a real-valued linear map on  $d \times d$  Hermitian matrices, which can be freely identified with a (not necessarily positive or normalized) Hermitian operator  $\rho \in i\mathfrak{u}(d)$ ,

$$\omega(A) = \text{Tr}(\rho A) \in \mathbb{R}. \quad (4.4.14)$$

The coadjoint action of a unitary  $U \in \mathbf{U}(d)$  on  $\omega \in i\mathfrak{u}(d)^*$  then corresponds to conjugation of  $\rho$ :

$$\mathrm{Ad}^*(U)(\rho)(A) = \omega(U^{-1}AU) = \mathrm{Tr}(U\rho U^{-1}A). \quad (4.4.15)$$

Furthermore, the maximal abelian subalgebra  $i\mathfrak{t} \cong \mathbb{R}^d$  is identified with the set of diagonal  $d \times d$  matrices with real entries, and the positive Weyl chamber those matrices with sorted entries along the diagonal. Therefore, if  $\omega$  is identified with  $\mathrm{Tr}(\rho \cdot)$  as above, then  $\omega_+$  (satisfying  $\omega = \mathrm{Ad}^*(U)(\omega_+)$ ) is identified by

$$\omega_+(A) = \mathrm{Tr}(\mathrm{diag}(s)A) \quad (4.4.16)$$

where  $s = \mathrm{spec}(\rho)$  is the vector of sorted eigenvalues of  $\rho$ .

Now consider the representation  $\Phi : \mathbf{U}(d) \rightarrow \mathbf{U}(dd')$  of  $\mathbf{U}(d)$  defined by

$$\Phi(U) = U \otimes \mathbb{1}_{d'}. \quad (4.4.17)$$

where  $d' \geq d$ . Then the moment map of a ray  $\psi \in \mathbb{P}(\mathbb{C}^d \otimes \mathbb{C}^{d'})$  then satisfies for all  $A \in i\mathfrak{u}(d)$ ,

$$\Omega_\Phi(\psi)(A) = \mathrm{Tr}_{d,d'}(P_\psi(A \otimes \mathbb{1}_d)) = \mathrm{Tr}_d(\sigma A). \quad (4.4.18)$$

where  $\sigma \in \mathcal{S}(\mathbb{C}^d)$  is the  $d$ -dimensional reduced density matrix  $\sigma = \mathrm{Tr}_{d'}(P_\psi)$ . Meanwhile the  $\omega$ -capacity of the ray  $\psi$  satisfies

$$\mathrm{cap}_\Phi^\omega(\psi)^2 = \inf_{M \in \mathrm{GL}(d, \mathbb{C})} e^{2\omega_+(\alpha_-(MU))} \mathrm{Tr}(M\sigma M^*), \quad (4.4.19)$$

$$= \inf_{M \in \mathrm{GL}(d, \mathbb{C})} e^{2\langle s, \alpha_-(M) \rangle} \mathrm{Tr}(M(U^{-1}\sigma U)M^*). \quad (4.4.20)$$

The details of this optimization have been worked out in [FW20] and [BCV21] to obtain the rate function defined by Eq. (4.4.4) of the form

$$I_\psi(\omega) = \sum_{j=1}^d s_j \log s_j - (s_j - s_{j-1}) \log \mathrm{lpm}_j(U^{-1}\sigma U), \quad (4.4.21)$$

where  $\mathrm{lpm}_j$  is the  $j$ th leading principal minor, whenever the spectra  $s \in \mathbb{R}^d$  of  $\rho$  is normalized and positive (so that  $\rho$  is a density operator), and otherwise  $I_\psi(\omega) = \infty$ . Furthermore, as  $\Omega_\Phi(\psi)(A) = \mathrm{Tr}(\sigma A)$  and  $\omega(A) = \mathrm{Tr}(\rho A)$ , we have

$$I_\psi(\omega) = 0 \iff \sigma = \rho. \quad (4.4.22)$$

In Chapter 6, specifically Corollary 6.5.11, we rediscover this rate function and refer to it as the *Keyl divergence* and use the notation  $K(\rho \parallel \sigma)$  in place of  $I_\psi(\omega)$ .

**Remark 4.4.6.** As a special case, one can consider applying the moment map estimation scheme from Definition 4.4.1 to  $n$  copies of the uniform state,

$$\sigma_n = \left( \frac{\mathbb{1}_{\mathcal{H}}}{d} \right)^{\otimes n}. \quad (4.4.23)$$

where  $d = \dim(\mathcal{H})$ . As  $\mathbb{1}_{\mathcal{H}}$  is  $K$ -invariant,  $\Phi(k^{-1})\mathbb{1}_{\mathcal{H}}\Phi(k) = \mathbb{1}_{\mathcal{H}}$ , the resulting probability measure over  $i\mathfrak{k}^*$  is constant along coadjoint orbits. Integrating over  $k \in K$  then yields a discrete measure  $\nu \rightarrow \Sigma(\Lambda_+) \rightarrow [0, 1]$  over highest weights where  $\lambda \in \Lambda_+$  has the probability

$$\nu(\{\lambda\}) = \frac{d_\lambda \text{Tr}(\Pi_{\Phi^{\otimes n}}^\lambda)}{d^n} = \frac{\dim(\mathcal{H}_\lambda) \dim(\mathcal{M}_\lambda^{\mathcal{H}^{\otimes n}})}{\dim(\mathcal{H})^n}. \quad (4.4.24)$$

Therefore, the large deviations results of [Theorem 4.4.2](#) and [Remark 4.4.3](#) recover, as special cases, the result of Cegła, Lewis, and Raggio [[CLR88](#)] on representations of  $K = \text{SU}(2)$  or more generally the result of Duffield [[Duf90](#)].

## 4.4.2 General definition

**Definition 4.4.7.** Let  $(X, \Sigma(X))$  be a standard Borel space and  $(\mathcal{S}(\mathcal{H}), \Sigma(\mathcal{S}(\mathcal{H})))$  the standard Borel space of density operators on a finite-dimensional complex Hilbert space  $\mathcal{H}$ . A measurable function  $f : \mathcal{S}(\mathcal{H}) \rightarrow X$  is called a **property** and elements  $x \in X$  are called **property values**.

**Example 4.4.8.** For example, a Boolean-valued property,  $p : \mathcal{S} \rightarrow \{T, F\}$ , of quantum states is just a proposition about quantum states.

Given some property  $f : \mathcal{S}(\mathcal{H}) \rightarrow X$  of quantum states, the purpose of an  $f$ -estimation scheme is to produce, for each state  $\rho$ , an estimate for the value of  $f(\rho)$ . Following the framework of Keyl [[Key06](#)], an *estimation scheme* will be modeled by a sequence of quantum measurements indexed by a positive integer  $n$  which indicates the number of copies of the underlying state  $\rho$  on which the measurement is performed. Formally, for each  $n$ , the associated measurement is described by a positive operator-valued measure (POVM) of the form

$$E_n : \Sigma(X) \rightarrow \mathcal{B}(\mathcal{H}^{\otimes n}). \quad (4.4.25)$$

Specifically,  $E_n$  is  $\sigma$ -additive set function from the Borel  $\sigma$ -algebra  $\Sigma(X)$  for  $X$  to the space of bounded operators on  $\mathcal{H}^{\otimes n}$  (denoted  $\mathcal{B}(\mathcal{H}^{\otimes n})$ ) such that  $E_n(\emptyset) = 0$ ,  $E_n(X) = \mathbb{1}_{\mathcal{H}^{\otimes n}}$ , and for all  $\Delta \in \Sigma(X)$  we have  $E_n(\Delta) \geq 0$ .

For a given positive integer  $n$ , measurement  $E_n : \Sigma(X) \rightarrow \mathcal{B}(\mathcal{H}^{\otimes n})$  and state  $\rho \in \mathcal{S}$ , the probability for obtaining a measurement outcome  $x$  in the region  $\Delta \in \Sigma(X)$  after performing the measurement  $E_n$  on  $\rho^{\otimes n}$  is expressed as

$$\xi_n^\rho(\Delta) := \xi_n(\Delta|\rho) := \text{Tr}(E_n(\Delta)\rho^{\otimes n}). \quad (4.4.26)$$

In order for the outcome of such a measurement to serve as an estimate for the property value of  $f(\rho)$ , at the very least, one should demand that the associated sequence of probability measures  $(\xi_n^\rho)_{n \in \mathbb{N}}$ , becomes increasingly concentrated around the true property value  $f(\rho) \in X$  as  $n \rightarrow \infty$ . By concentration around the value  $f(\rho) \in X$ , we consider the Dirac measure  $\delta_{f(\rho)} : \Sigma(X) \rightarrow [0, 1]$ , localized at  $f(\rho) \in X$ , defined by

$$\delta_{f(\rho)}(\Delta) = \begin{cases} 1 & x \in \Delta, \\ 0 & x \notin \Delta. \end{cases} \quad (4.4.27)$$

By convergence in this instance, we demand that for all bounded and continuous functions,  $g : X \rightarrow \mathbb{R}$ , the expected value of  $g$ , taken with respect to the state-dependent measure  $\xi_n^\rho$  converges to the value of  $g(f(\rho))$ :

$$\lim_{n \rightarrow \infty} \int_X g \xi_n^\rho = g(f(\rho)). \quad (4.4.28)$$

In other words, we demand that the sequence  $(\xi_n^\rho)_{n \in \mathbb{N}}$  of probability measures on  $X$  *converges weakly* (see [Section 2.1](#) for a definition of weak convergence) to the aforementioned Dirac measure,  $\delta_{f(\rho)}$  on  $X$ .

**Example 4.4.9.** For the sake of concreteness, consider briefly the task of estimating the spectrum of a qutrit quantum state,  $\rho \in \mathcal{S}(\mathbb{C}^3)$ . In this case, the property under consideration is the function which maps a qutrit density operator  $\rho$  to its three non-negative eigenvalues,  $\text{spec}(\rho) = (\lambda_1, \lambda_2, \lambda_3)$ , which may be assumed sorted and summing to one. The space of all possible property values is therefore the set  $X = \Sigma_3^\downarrow$  where

$$\Sigma_3^\downarrow := \{(\lambda_1, \lambda_2, \lambda_3) \in \mathbb{R}^3 \mid \lambda_1 \geq \lambda_2 \geq \lambda_3 \geq 0, \lambda_1 + \lambda_2 + \lambda_3 = 1\}. \quad (4.4.29)$$

Using the spectral estimation scheme proposed by Keyl and Werner [[KW01](#)], to be considered in detail in [Section 4.4.4](#), one can perform a collective measurement on  $n$  copies of the state  $\rho$  and extract an estimate  $(\lambda'_1, \lambda'_2, \lambda'_3) \in \Sigma_3^\downarrow$  for the spectrum of  $\rho$  with probability measure  $\xi_n^\rho((\lambda'_1, \lambda'_2, \lambda'_3))$ . As the number  $n$  of copies of  $\rho$  increases, the probability measure for estimates should converge to the Dirac measure concentrated at the true spectrum of  $\rho$ , as illustrated by [Fig. 4.1](#).

An alternative way to formalize this condition without making reference to functions on  $X$  is to require, for every  $\epsilon > 0$ , that the probability of obtaining an estimate,  $x \in X$ , which is a distance of  $\epsilon$  (or greater) from the true value,  $f(\rho) \in X$ , tends to zero as  $n \rightarrow \infty$ . Stated more formally, if

$$B_\epsilon(f(\rho)) := \{x \in X \mid d(x, f(\rho)) < \epsilon\} \quad (4.4.30)$$

is the open ball of radius  $\epsilon$  around  $f(\rho)$ , then one should demand

$$\lim_{n \rightarrow \infty} \xi_n^\rho(X \setminus B_\epsilon(f(\rho))) = 0. \quad (4.4.31)$$

In other words, for any distance  $\epsilon > 0$  the probability that the obtained estimate  $x$  is within a distance of  $\epsilon$  of the true value  $f(\rho)$  can be made arbitrarily close to unity by taking  $n$  sufficiently large.

**Remark 4.4.10.** The two conditions expressed by [Eq. \(4.4.28\)](#) and [Eq. \(4.4.31\)](#) are actually equivalent. This equivalence follows from an application of the Portmanteau theorem (see [Theorem 2.1.11](#)) together with the fact that the complement,  $X \setminus B_\epsilon(f(\rho))$ , of the open ball  $B_\epsilon(f(\rho))$  is both closed and excludes  $f(\rho)$  and furthermore, every closed subset  $C \subseteq X$  which excludes  $f(\rho)$  is a subset of the complement  $X \setminus B_\epsilon(f(\rho))$  for some sufficiently small  $\epsilon > 0$ .

This brings us to our formal definition of an  $f$ -estimation scheme.

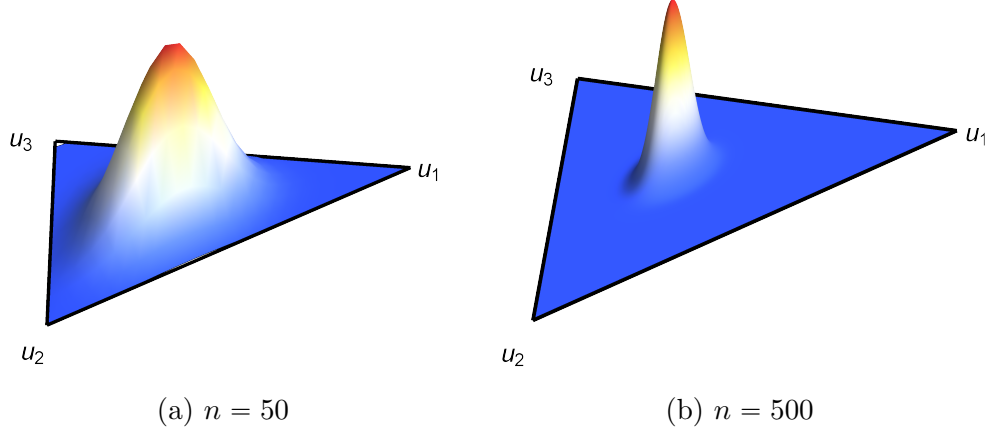


Figure 4.1: A visualization of the probability measure,  $\xi_n^\rho((\lambda_1, \lambda_2, \lambda_3)|\rho)$ , of estimates for the spectrum of a qutrit quantum state,  $\rho$ , with spectrum  $\text{spec}(\rho) = (0.6, 0.3, 0.1)$  when the spectral estimation scheme proposed by Keyl and Werner's [KW01] is applied to  $n$  copies of  $\rho$ . The domain in the figures corresponds to the set of all possible sorted spectra  $(\lambda_1, \lambda_2, \lambda_3)$  of a qutrit (Eq. (4.4.29)) and has extremal vertices given by  $u_1 = (1, 0, 0)$  (pure state),  $u_2 = (1/2, 1/2, 0)$  and  $u_3 = (1/3, 1/3, 1/3)$  (maximally mixed). This figure is a reproduction of [KW01, Fig. 1], but for different values of  $n$ .

**Definition 4.4.11.** Let  $(E_n : \Sigma(X) \rightarrow \mathcal{B}(\mathcal{H}^{\otimes n}))_{n \in \mathbb{N}}$  be a sequence of POVMs and let  $f : \mathcal{S}(\mathcal{H}) \rightarrow X$  be a property of quantum states. If, for every state,  $\rho \in \mathcal{S}(\mathcal{H})$ , and for every bounded, continuous function  $g : X \rightarrow \mathbb{R}$ , we have

$$\lim_{n \rightarrow \infty} \int_{x \in X} g(x) \text{Tr}[dE_n(x) \rho^{\otimes n}] = g(f(\rho)), \quad (4.4.32)$$

then the sequence,  $(E_n)_{n \in \mathbb{N}}$ , of POVMs is said to be an ***f-estimation scheme***.

**Remark 4.4.12.** While our proposed definition for an  $f$ -estimation scheme is directly inspired by the definition of a property estimation scheme from Keyl's paper [Key06] on quantum state estimation, our usage of the terminology is slightly different. Strictly speaking, in [Key06], any sequence of POVMs of the form  $(E_n : \Sigma(X) \rightarrow \mathcal{B}(\mathcal{H}^{\otimes n}))_{n \in \mathbb{N}}$  is called an estimation scheme, and those sequences of POVMs which satisfy Eq. (4.4.32) are said to be *consistent* with  $f$ .

**Definition 4.4.13.** An  $f$ -estimation scheme  $(E_n : \Sigma(X) \rightarrow \mathcal{B}(\mathcal{H}^{\otimes n}))_{n \in \mathbb{N}}$  is said to be

- (i) ***discrete*** if, for each  $n$ , there exists a countably infinite subset  $D_n \subseteq X$  such that

$$\forall \Delta \in \Sigma(X) : E_n(\Delta) = \sum_{x \in D_n \cap \Delta} E_n(\{x\}), \quad (4.4.33)$$

- (ii) ***finite*** if it is discrete and moreover the set  $D_n$  has finite cardinality,

$$\forall n \in \mathbb{N} : |D_n| < \infty, \quad (4.4.34)$$

(iii) **projective** if for each  $n$ ,  $E_n$  is a projective-valued measure, i.e.

$$\forall \Delta_1, \Delta_2 \in \Sigma(X) : E_n(\Delta_1)E_n(\Delta_2) = E_n(\Delta_1 \cap \Delta_2). \quad (4.4.35)$$

In addition to the measure-theoretic structure of estimation schemes, they also tend to exhibit certain symmetries. In particular, for each integer  $n \in \mathbb{N}$ , the tensor-permutation representation,  $T_n : S_n \rightarrow \text{GL}(\mathcal{H}^{\otimes n})$ , of the symmetric group acting on the  $n$ -fold tensor product  $\mathcal{H}^{\otimes n}$  (defined in [Example 2.4.2](#)) induces group actions of  $S_n$  on  $\mathcal{B}(\mathcal{H}^{\otimes n})$  and thus on estimation schemes themselves. Similarly, if  $\Phi : G \rightarrow \text{GL}(\mathcal{H})$  is a representation of a group  $G$  on  $\mathcal{H}$ , then one can consider the tensor-power representation of  $G$  on  $\mathcal{H}^{\otimes n}$  and its induced action on estimation schemes.

**Definition 4.4.14.** Let  $\Phi : G \rightarrow \text{GL}(\mathcal{H})$  be a group representation of  $G$  on  $\mathcal{H}$  and, for each  $n \in \mathbb{N}$ , let  $T_n : S_n \rightarrow \text{GL}(\mathcal{H}^{\otimes n})$  be the tensor-permutation representation of  $S_n$  on  $\mathcal{H}^{\otimes n}$ . An  $f$ -estimation scheme  $(E_n : \Sigma(X) \rightarrow \mathcal{B}(\mathcal{H}^{\otimes n}))_{n \in \mathbb{N}}$  is said to be

(i)  **$G$ -covariant** if there exists a continuous map  $\alpha : G \times X \rightarrow X :: (g, x) \mapsto \alpha_g(x)$  such that for all  $g \in G$ ,  $\Delta \in \Sigma(X)$  and  $n \in \mathbb{N}$ ,

$$\Phi^{\otimes n}(g)E_n(\Delta)(\Phi^{\otimes n}(g))^* = E_n(\alpha_g(\Delta)), \quad (4.4.36)$$

where  $\alpha_g(\Delta) = \{\alpha_g(x) \in X \mid x \in \Delta\}$ ,

(ii)  **$G$ -invariant** if it is  $G$ -covariant and where the map  $\alpha$  satisfies  $\alpha_g(x) = x$  for all  $g \in G$  and  $x \in X$ ,

(iii)  **$S_n$ -invariant** if for all  $n \in \mathbb{N}$ ,  $\sigma \in S_n$  and  $\Delta \in \Sigma(X)$ ,

$$T_n(\sigma)E_n(\Delta)T_n^*(\sigma) = E_n(\Delta). \quad (4.4.37)$$

### 4.4.3 Derived estimation schemes

In [Section 4.4.4](#), we will see that almost all examples of estimation schemes are derived from manipulating another estimation scheme.

Our first result concerning the compositionality of estimation schemes is concerned with the case of transforming an  $f$ -estimation scheme into an  $f'$ -estimation scheme when the property  $f' : \mathcal{S}(\mathcal{H}) \rightarrow X'$  is a function of the property  $f : \mathcal{S}(\mathcal{H}) \rightarrow X$ . In other words, if there exists a transition function  $g : X \rightarrow X'$  such that  $f' = g \circ f$ , then the  $f$ -estimation scheme  $(E_n)_{n \in \mathbb{N}}$  which, for each state  $\rho \in \mathcal{S}$ , produces an estimate for the value of  $f(\rho)$ , should be capable of producing an estimate for the value of  $f'(\rho) = g(f(\rho))$ . Of course, in order to ensure the appropriate convergence, it suffices to assume that  $g$  is continuous.

**Lemma 4.4.15.** *Let  $f : \mathcal{S}(\mathcal{H}) \rightarrow X$  be a property and let  $(E_n : \Sigma(X) \rightarrow \mathcal{B}(\mathcal{H}^{\otimes n}))_{n \in \mathbb{N}}$  be an  $f$ -estimation scheme. If  $g : X \rightarrow X'$  is a continuous, measurable function, then the **pushforward estimation scheme** of  $E_n$  by  $g$ , denoted by*

$$(g_*E_n : \Sigma(X') \rightarrow \mathcal{B}(\mathcal{H}^{\otimes n}))_{n \in \mathbb{N}} \quad (4.4.38)$$

and defined for all  $\Delta' \in \Sigma(X')$  and  $n \in \mathbb{N}$  by

$$(g_*E_n)(\Delta') = E_n(g^{-1}(\Delta')), \quad (4.4.39)$$

is a  $(g \circ f)$ -estimation scheme.

*Proof.* The proof follows directly from [Lemma 2.1.12](#).  $\square$

One might wonder what can be said about the case where the transition function  $g : X \rightarrow X'$  happens to be discontinuous. A common instance of such a discontinuous function<sup>4</sup> is the indicator function  $\text{in}_R : X \rightarrow \{F, T\}$  which determines whether  $x \in X$  belongs to the subset  $R \subseteq X$ :

$$\text{in}_R(x) := \begin{cases} T & x \in R, \\ F & x \notin R. \end{cases} \quad (4.4.40)$$

The composition,  $\text{in}_R \circ f$ , then corresponds to the proposition that evaluates whether or not the property value  $f(\rho)$  lies in  $R \subseteq X$ . A natural question arises, given an  $f$ -estimation scheme  $(E_n)_{n \in \mathbb{N}}$ , how can one construct an  $(\text{in}_R \circ f)$ -estimation scheme? One might think that if the region  $R$  itself is already a Borel set, i.e.  $R \in \Sigma(X)$ , then the two-outcome POVM defined by  $\{E_n(R), E_n(X \setminus R)\}$ , taken for all  $n$ , would serve as an  $(\text{in}_R \circ f)$ -estimation scheme. Unfortunately, as the next example demonstrates, this is not always the case.

**Example 4.4.16.** Let  $\mathcal{H} \cong \mathbb{C}^2$  be the Hilbert space associated to a qubit, let  $P_\psi = |\psi\rangle\langle\psi|$  be a rank one projection operator, let  $P_{-\psi} = \mathbb{1} - P_\psi$ , and let  $f : \mathcal{S}(\mathbb{C}^2) \rightarrow [0, 1]$  be the property which corresponds to the probability of obtaining the outcome  $P_\psi$  when applying the projective measurement  $\{P_\psi, P_{-\psi}\}$ , i.e.,  $f(\rho) = \text{Tr}(\rho P_\psi)$ . An estimation scheme,  $E_n : \Sigma([0, 1]) \rightarrow \mathcal{B}((\mathbb{C}^2)^{\otimes n})$ , for  $p = \text{Tr}(\rho P_\psi)$  emerges from projectively measuring  $\{P_\psi, P_{-\psi}\}$  a total of  $n \in \mathbb{N}$  times and recording the number  $k \in \{0, \dots, n\}$  of instances of the first outcome. More formally, for any Borel subset  $\Delta \in \Sigma([0, 1])$  and integer  $n \in \mathbb{N}$ , let

$$E_n(\Delta) = \sum_{\substack{0 \leq k \leq n \\ k/n \in \Delta}} P_{(k)} \quad (4.4.41)$$

where  $k \in \{0, 1, \dots, n\}$  and  $P_{(k)} \in \text{End}((\mathbb{C}^2)^{\otimes n})$  is the projection operator

$$P_{(k)} = \frac{1}{n!} \binom{n}{k} \sum_{\Phi \in \mathcal{S}_n} T_\Phi(P_\psi^{\otimes k} \otimes P_{-\psi}^{\otimes(n-k)}) T_{\Phi^{-1}}. \quad (4.4.42)$$

It is not too difficult to see that this forms an estimation scheme for  $p = f(\rho) = \text{Tr}(\rho P_\psi)$  in the sense that for all states  $\rho$ ,

$$\text{Tr}(\rho P_\psi) = \lim_{n \rightarrow \infty} \int_0^1 p \text{Tr}(dE_n(p) \rho^{\otimes n}). \quad (4.4.43)$$

Now consider the coarse-grained property,

$$f' := \text{in}_{\{\frac{1}{2}\}} \circ f : \mathcal{S}(\mathbb{C}^2) \rightarrow \{F, T\}, \quad (4.4.44)$$

which decides whether or not  $\text{Tr}(\rho P_\psi) = \frac{1}{2}$ . Using the  $f$ -estimation scheme  $(E_n)_{n \in \mathbb{N}}$  defined above, one can consider a corresponding coarse-grained POVM

$$E'_n : \Sigma(\{T, F\}) \rightarrow \mathcal{B}((\mathbb{C}^2)^{\otimes n}) \quad (4.4.45)$$

---

<sup>4</sup>Assuming that the topology on  $X$  is such that  $\text{in}_R$  is not already continuous.



defined according to

$$E'_n(\{T\}) = E_n(\{\frac{1}{2}\}) = \begin{cases} P_{(\frac{n}{2})} & n \text{ is even,} \\ 0 & n \text{ is odd.} \end{cases} \quad (4.4.46)$$

Here it is not too difficult to see that even if the state  $\sigma$  satisfies  $\text{Tr}(\sigma P_\psi) = \frac{1}{2}$  and thus  $f'(\sigma) = T$ , we observe

$$\limsup_{n \rightarrow \infty} \text{Tr}[E'_n(\{F\})\sigma^{\otimes n}] = 1, \quad (4.4.47)$$

because it is impossible to produce an estimate for  $\text{Tr}(\sigma P_\psi)$  of  $\frac{1}{2}$  after an odd number of trials and therefore  $E'_n$  is *not* an  $f'$ -estimation scheme.

Despite the apparent obstacle presented by [Example 4.4.16](#) for elegantly constructing  $\text{in}_R \circ f$ -estimation schemes from  $f$ -estimation schemes, the following result proves it is always possible using an alternative method.

**Lemma 4.4.17.** *Let  $(E_n)_{n \in \mathbb{N}}$  be an  $f$ -estimation scheme for the property  $f : \mathcal{S}(\mathcal{H}) \rightarrow X$  and let  $R \in \Sigma(X)$  be closed. Then there exists an  $(\text{in}_R \circ f)$ -estimation scheme, denoted by  $(E_n^R)_{n \in \mathbb{N}}$ , of the form*

$$E_n^R(\{T\}) = E_n(\Delta_n), \quad \text{and} \quad E_n^R(\{F\}) = E_n(X \setminus \Delta_n), \quad (4.4.48)$$

where  $(\Delta_n \in \Sigma(X))_{n \in \mathbb{N}}$  is a sequence of closed, nested subsets converging to  $R$  in the sense that  $\bigcap_{n \in \mathbb{N}} \Delta_n = R$ .

*Proof.* The proof is analogous to the proof of [Lemma 2.1.18](#) and will not be repeated here. In the end, for  $n \in \mathbb{N}$ , the subset  $\Delta_n$  is the closure of  $B_{q(n)}(R)$  where, for  $\epsilon > 0$ ,

$$B_\epsilon(R) := \{x \in X \mid \exists r \in R, d(x, r) < \epsilon\}, \quad (4.4.49)$$

and where  $q(n)$  is a non-increasing sequence with limit  $\lim_{n \rightarrow \infty} q(n) = 0$ . □

The following result is similar to [Lemma 4.4.15](#) in that it constructs a new estimation scheme from an old estimation scheme, but dissimilar in that it modifies the state space,  $\mathcal{S}(\mathcal{H})$ , of the function  $f : \mathcal{S}(\mathcal{H}) \rightarrow X$  instead of the property space,  $X$ .

**Proposition 4.4.18.** *Let  $f : \mathcal{S}(\mathcal{H}) \rightarrow X$  be a property and let  $(E_n : \Sigma(X) \rightarrow \mathcal{B}(\mathcal{H}^{\otimes n}))_{n \in \mathbb{N}}$  be an  $f$ -estimation scheme. Let  $\mathcal{C} : \mathcal{B}(\mathcal{H}') \rightarrow \mathcal{B}(\mathcal{H})$  be a completely positive, trace preserving map with Kraus decomposition*

$$\mathcal{C}(\rho) = \sum_k C_k \rho C_k^* \quad (4.4.50)$$

where  $\{C_k : \mathcal{H}' \rightarrow \mathcal{H}\}_k$  are linear transformations satisfying  $\sum_k C_k^* C_k = \mathbb{1}_{\mathcal{H}'}$ . Then the sequence of POVMs  $(E'_n : \Sigma(X) \rightarrow \mathcal{B}((\mathcal{H}')^{\otimes n}))_{n \in \mathbb{N}}$  defined for  $\Delta \in \Sigma(X)$  by

$$E'_n(\Delta) = \sum_{k_1, \dots, k_n} (C_{k_1} \otimes \dots \otimes C_{k_n})^* E_n(\Delta) (C_{k_1} \otimes \dots \otimes C_{k_n}) \quad (4.4.51)$$

is a  $(f \circ \mathcal{C})$ -estimation scheme where  $f \circ \mathcal{C} : \mathcal{S}(\mathcal{H}') \rightarrow X$ .



The intuition behind [Proposition 4.4.18](#) is simply that if one can estimate the value of  $f(\sigma) \in X$  where the state  $\sigma$  is the result of sending the state  $\rho$  through the channel  $\mathcal{C}$  such that  $\sigma = \mathcal{C}(\rho)$ , then one is implicitly estimating the value of  $(f \circ \mathcal{C})(\rho)$ . Our primary application of [Proposition 4.4.18](#) will be to the estimation of marginal or local properties of a composite quantum state, i.e., where the channel  $\mathcal{C}$  is simply a partial trace operation.

**Example 4.4.19.** Suppose the Hilbert space  $\mathcal{H}$  is the tensor product of two Hilbert spaces, i.e.  $\mathcal{H} \cong \mathcal{A} \otimes \mathcal{B}$ , and suppose one is interested in a property  $g : \mathcal{S}(\mathcal{A} \otimes \mathcal{B}) \rightarrow X$  which is local to subsystem  $\mathcal{A}$ , meaning

$$g(\rho_{AB}) = f(\text{Tr}_B(\rho_{AB})), \quad (4.4.52)$$

for some property  $f : \mathcal{S}(\mathcal{A}) \rightarrow X$ . For instance,  $f$  could be the expectation value of a local observable where

$$g(\rho_{AB}) = \text{Tr}_{AB}((O_A \otimes \mathbb{1}_B)\rho_{AB}) = \text{Tr}_A(O_A \rho_A) = f(\rho_A). \quad (4.4.53)$$

Any  $f$ -estimation scheme  $(E_n^A : \Sigma(X) \rightarrow \mathcal{B}(\mathcal{H}_A^{\otimes n}))_{n \in \mathbb{N}}$  can be *lifted* to a  $g$ -estimation scheme  $(E_n^{AB} : \Sigma(X) \rightarrow \mathcal{B}((\mathcal{H}_A \otimes \mathcal{H}_B)^{\otimes n}))_{n \in \mathbb{N}}$  by defining, for each  $n \in \mathbb{N}$ ,

$$E_n^{AB}(\Delta) := E_n^A(\Delta) \otimes \mathbb{1}_B^{\otimes n}. \quad (4.4.54)$$

In addition to [Lemma 4.4.15](#), there is another way to reuse estimation schemes by exploiting the fact that  $\text{Tr}(E_k(\Delta)\rho^{\otimes k})$  can be viewed as a polynomial in  $\rho$  of degree  $k$  or as a linear function in  $\rho^{\otimes k}$ .

**Proposition 4.4.20.** *Let  $k \in \mathbb{N}$  be a positive integer, let  $f : \mathcal{S}(\mathcal{H}^{\otimes k}) \rightarrow X$  be a property and let  $(E_n : \Sigma(X) \rightarrow \mathcal{B}((\mathcal{H}^{\otimes k})^{\otimes n}))_{n \in \mathbb{N}}$  be an  $f$ -estimation scheme. If  $g : \mathcal{S}(\mathcal{H}) \rightarrow \mathcal{S}(\mathcal{H}^{\otimes k})$  is the mapping sending  $\rho \in \mathcal{S}(\mathcal{H})$  to  $\rho^{\otimes k} \in \mathcal{S}(\mathcal{H}^{\otimes k})$ , then the sequence of POVMs  $(E'_n : \Sigma(X) \rightarrow \mathcal{B}(\mathcal{H}^{\otimes n}))_{n \in \mathbb{N}}$  defined for  $\Delta \in \Sigma(X)$  by*

$$E'_n(\Delta) = E_{\lfloor \frac{n}{k} \rfloor}(\Delta) \otimes \mathbb{1}_{\mathcal{H}}^{\otimes (n \bmod k)}, \quad (4.4.55)$$

is a  $(f \circ g)$ -estimation scheme.

The intended use case for [Proposition 4.4.20](#) is for the estimation of properties which can be seen as polynomial functions in the coefficients of the underlying states, otherwise known as *multi-copy observables* [[Ver+23](#)].

**Example 4.4.21.** Consider, for instance, the purity  $p(\rho) = \text{Tr}_{\mathcal{H}}(\rho^2)$  of a density operator  $\rho \in \mathcal{H}$ . It is well-known that the purity can be equivalently expressed as

$$p(\rho) = \text{Tr}_{\mathcal{H}}(\rho^2) = \text{Tr}_{\mathcal{H}^{\otimes 2}}(X_{\text{swap}}(\rho \otimes \rho)), \quad (4.4.56)$$

where  $X_{\text{swap}} = T_2((12)) \in \mathcal{H}^{\otimes 2}$  is the unitary and Hermitian operator which permutes the two tensor factors of  $\mathcal{H}^{\otimes 2}$ . Now consider the property  $f : \mathcal{S}(\mathcal{H}^{\otimes 2}) \rightarrow [-1, +1]$ , defined for  $\sigma \in \mathcal{S}(\mathcal{H}^{\otimes 2})$  as the expectation value

$$f(\sigma) = \text{Tr}(X_{\text{swap}}\sigma). \quad (4.4.57)$$

Given any scheme for estimating the expectation value of  $f(\sigma)$  of some unknown state  $\sigma \in \mathcal{S}(\mathcal{H}^{\otimes 2})$ , one could then estimate the purity  $p(\rho)$  of  $\rho$  using the correspondence  $f(\rho^{\otimes 2}) = p(\rho)$ . [Proposition 4.4.20](#) provides such a recipe for performing purity estimation by setting  $k = 2$  and viewing  $n$  copies of the state  $\rho$  as  $\lfloor \frac{n}{2} \rfloor$  copies of the state  $\rho^{\otimes 2}$  (with potentially one copy of  $\rho$  leftover) to be used to extract an estimate of  $f(\rho^{\otimes 2})$ .

Next consider the scenario where one has access to two estimation schemes,  $(E_n^1)_{n \in \mathbb{N}}$  for the property  $f_1 : \mathcal{S}(\mathcal{H}) \rightarrow X_1$  and  $(E_n^2)_{n \in \mathbb{N}}$  for the property  $f_2 : \mathcal{S}(\mathcal{H}) \rightarrow X_2$ . How does one estimate the composite property,  $(f_1, f_2) : \mathcal{S}(\mathcal{H}) \rightarrow X_1 \times X_2$ ? If the two estimation schemes commute, then we obtain the following result.

**Lemma 4.4.22** (Commuting Estimation Schemes). *If  $(E_n^1)_{n \in \mathbb{N}}$  is an  $f_1$ -estimation scheme and  $(E_n^2)_{n \in \mathbb{N}}$  is an  $f_2$ -estimation scheme, and for each  $n$ ,  $E_n^1$  and  $E_n^2$  are commuting, i.e.  $\forall \Delta_1 \in \Sigma(X_1), \Delta_2 \in \Sigma(X_2)$*

$$[E_n^1(\Delta_1), E_n^2(\Delta_2)] = 0, \quad (4.4.58)$$

then  $(E_n : \Sigma(X_1 \times X_2) \rightarrow \mathcal{B}(\mathcal{H}^{\otimes n}))_{n \in \mathbb{N}}$ , defined on  $\Delta_1 \times \Delta_2 \in \Sigma(X_1 \times X_2)$  as

$$E_n(\Delta_1 \times \Delta_2) := E_n^1(\Delta_1)E_n^2(\Delta_2) = E_n^2(\Delta_2)E_n^1(\Delta_1)$$

is an  $(f_1, f_2)$ -estimation scheme.

*Proof.* The proof makes use of the upper bounds

$$E_n(\Delta_1 \times \Delta_2) \leq E_n(\Delta_1 \times X_2) = E_n^1(\Delta_1), \quad \text{and} \quad (4.4.59)$$

$$E_n(\Delta_1 \times \Delta_2) \leq E_n(X_1 \times \Delta_2) = E_n^2(\Delta_2). \quad (4.4.60)$$

Therefore, for all  $\rho$ ,

$$\text{Tr}(E_n(\Delta_1 \times \Delta_2)\rho^{\otimes n}) \leq \min_{i \in \{1,2\}} \text{Tr}(E_n^i(\Delta_i)\rho^{\otimes n}). \quad (4.4.61)$$

Consequently, if  $C_1 \times C_2 \in \Sigma(X_1 \times X_2)$  is closed and *excludes* the composite property value  $(f_1, f_2)(\rho) \in X_1 \times X_2$ , and thus either  $f_1(\rho) \notin C_1$  or  $f_2(\rho) \notin C_2$  (or both), then at least one of the upper bounds provided above will tend to zero as  $n \rightarrow \infty$  which means

$$\limsup_{n \rightarrow \infty} \text{Tr}(E_n(C_1 \times C_2)\rho^{\otimes n}) = 0. \quad (4.4.62)$$

On the other hand, if  $C_1 \times C_2 \in \Sigma(X_1 \times X_2)$  is closed and *contains* the composite property value  $(f_1, f_2)(\rho) \in X_1 \times X_2$ , then we expect for each  $i \in \{1, 2\}$ ,

$$\limsup_{n \rightarrow \infty} \text{Tr}(E_n^i(C_i)\rho^{\otimes n}) = 1, \quad (4.4.63)$$

As  $E_n^1$  and  $E_n^2$  commute and are positive-operator-valued, we have a kind of non-projective and commutative union bound [[Gao15](#)],

$$\begin{aligned} E_n(C_1 \times C_2) &= E_n^1(C_1)E_n^2(C_2), \\ &= \sqrt{E_n^2(C_2)}E_n^1(C_1)\sqrt{E_n^2(C_2)}, \\ &= E_n^2(C_2) - \sqrt{E_n^2(C_2)}E_n^1(X_1 \setminus C_1)\sqrt{E_n^2(C_2)}, \\ &= E_n^2(C_2) - \sqrt{E_n^1(X_1 \setminus C_1)}E_n^2(C_2)\sqrt{E_n^1(X_1 \setminus C_1)}. \\ &\geq E_n^2(C_2) - E_n^1(X_1 \setminus C_1). \end{aligned} \quad (4.4.64)$$

By Eq. (4.4.63),  $\liminf_{n \rightarrow \infty} \text{Tr}(E_n^1(X_1 \setminus C_1)) = 0$ , and therefore

$$\begin{aligned}
& \limsup_{n \rightarrow \infty} \text{Tr}(E_n(C_1 \times C_2)\rho^{\otimes n}) \\
& \geq \limsup_{n \rightarrow \infty} (\text{Tr}(E_n^2(C_2)\rho^{\otimes n}) - \text{Tr}(E_n^1(X_1 \setminus C_1)\rho^{\otimes n})) \\
& = \limsup_{n \rightarrow \infty} \text{Tr}(E_n^2(C_2)\rho^{\otimes n}) - \liminf_{n \rightarrow \infty} \text{Tr}(E_n^1(X_1 \setminus C_1)\rho^{\otimes n}) \\
& = 1 - 0 = 1.
\end{aligned} \tag{4.4.65}$$

Together, Eq. (4.4.62) and Eq. (4.4.65) imply (via condition (iii) of Theorem 2.1.11) that  $(E_n : \Sigma(X_1 \times X_2) \rightarrow \mathcal{B}(\mathcal{H}^{\otimes n}))_{n \in \mathbb{N}}$  as defined above is an estimation scheme for the composite property  $(f_1, f_2)$ .  $\square$

Now suppose the two estimation schemes,  $(E_n^1)_{n \in \mathbb{N}}$  for  $f_1 : \mathcal{S}(\mathcal{H}) \rightarrow X_1$  and  $(E_n^2)_{n \in \mathbb{N}}$  for  $f_2 : \mathcal{S}(\mathcal{H}) \rightarrow X_2$ , are not commuting. How, then, does one estimate the product property,  $f_1, f_2 : \mathcal{S}(\mathcal{H}) \rightarrow X_1 \times X_2$ ? Perhaps the simplest option is to partition the  $n$  independent and identically prepared copies of  $\rho$  into two disjoint blocks and use the first block to estimate the value of  $f_1(\rho)$  and the second block to estimate the value of  $f_2(\rho)$ .

**Lemma 4.4.23** (Non-Commuting). *For each  $n \in \mathbb{N}$ , let  $b_n = \lfloor \frac{n}{2} \rfloor$  and  $b'_n = n - b_n$ . If  $(E_n^1)_{n \in \mathbb{N}}$  is an  $f_1$ -estimation scheme and  $(E_n^2)_{n \in \mathbb{N}}$  is an  $f_2$ -estimation scheme, then*

$$E_n(\Delta_1 \times \Delta_2) := E_{b_n}^1(\Delta_1) \otimes E_{b'_n}^2(\Delta_2)$$

is an  $(f_1, f_2)$ -estimation scheme.

*Proof.* As the two estimation schemes have been made to act on disjoint blocks of  $\mathcal{H}^{\otimes n}$ , the two estimation schemes effectively commute with each other,

$$[E_{b_n}^1(\Delta_1) \otimes \mathbb{1}_{\mathcal{H}}^{\otimes b'_n}, \mathbb{1}_{\mathcal{H}}^{\otimes b_n} \otimes E_{b'_n}^2(\Delta_2)] = 0. \tag{4.4.66}$$

Therefore, the proof can be seen as a special case of Lemma 4.4.22. The essential difference however is to notice that (i) as  $n \rightarrow \infty$ , the two block sizes,  $b_n = \lfloor \frac{n}{2} \rfloor$  and  $b'_n = n - b_n$ , tend to infinity, and moreover, (ii) every integer is contained in the sequences  $(b_n)_{n \in \mathbb{N}}$  and  $(b'_n)_{n \in \mathbb{N}}$ . This secondary observation is needed to ensure that all of the limits still hold.  $\square$

#### 4.4.4 Example estimation schemes

In this section we endeavour to list a wide variety of estimation schemes for various properties of quantum states. While many of the following examples are examples of *moment map* estimation schemes from Section 4.4.1, some of them are not of this form.

**Example 4.4.24.** Consider a  $d$ -dimensional Hilbert space,  $\mathcal{H} \cong \mathbb{C}^d$ , and let  $X \in \text{End}(\mathcal{H})$  be a Hermitian operator,  $X^* = X$ . Now consider the property,  $\langle X \rangle : \mathcal{S}(\mathbb{C}^d) \rightarrow \mathbb{R}$ , assigning to each state its  $X$ -expectation value,

$$\rho \mapsto \langle X \rangle_\rho = \text{Tr}(\rho X). \tag{4.4.67}$$

Let the spectral decomposition of  $X$  be

$$X = \sum_{j=1}^d x_j P_j. \quad (4.4.68)$$

Now for each  $n \in \mathbb{N}$ , imagine performing the projective measurement,  $\{P_j\}_{j=1}^d$ ,  $n$  times and letting  $(\lambda_1, \dots, \lambda_d) \in \mathbb{N}^d$  record the number of outcomes of each type. Then the collective POVM,  $E_n^X : \Sigma(\mathbb{R}) \rightarrow \mathcal{B}(\mathcal{H}^{\otimes n})$ , defined by

$$E_n^X(\Delta) = \sum_{\substack{(\lambda_1, \dots, \lambda_d) \\ \frac{1}{n} \sum_j x_j \lambda_j \in \Delta}} \frac{1}{\lambda_1! \cdots \lambda_d!} \sum_{\pi \in S_n} T_n(\pi) \left( P_1^{\otimes \lambda_1} \otimes \cdots \otimes P_d^{\otimes \lambda_d} \right) T_n(\pi^{-1}), \quad (4.4.69)$$

is an  $\langle X \rangle$ -estimation scheme.

**Example 4.4.25.** Our next example is the spectral estimation scheme due to Keyl and Werner [KW01]. Let  $\text{spec} : \mathcal{S}(\mathbb{C}^d) \rightarrow \Delta_d^\downarrow$  be the function sending each  $d$ -dimensional density operator,  $\rho$ , to its sorted spectrum of non-negative eigenvalues summing to one. Note that the irreducible finite-dimensional representations of  $\text{SU}(d)$  are indexed by lists of  $d$  non-negative integers,  $(\lambda_1, \dots, \lambda_d) \in \mathbb{N}^d$ , which are non-increasing,  $\lambda_1 \geq \cdots \geq \lambda_d$ . Furthermore, the irreducible subrepresentations of the representation  $U \in \text{SU}(d)$  to  $U^{\otimes n}$  acting on  $(\mathbb{C}^d)^{\otimes n}$  are indexed by those  $\lambda$  whose entries total  $n$ , i.e.,  $\lambda_1 + \dots + \lambda_d = n$ . For each  $n \in \mathbb{N}$ , let  $E_n^{\text{spec}} : \Sigma(\Delta_d^\downarrow) \rightarrow \mathcal{B}((\mathbb{C}^d)^{\otimes n})$  be the POVM defined for all  $B \in \Sigma(\Delta_d^\downarrow)$  by

$$E_n^{\text{spec}}(B) = \sum_{\substack{\lambda \in \mathbb{Y}_d^n \\ \frac{\lambda}{n} \in B}} \text{Tr}(\tilde{\Pi}_{\mathcal{H}^{\otimes n}}^\lambda \rho^{\otimes n}), \quad (4.4.70)$$

where  $\tilde{\Pi}_{\mathcal{H}^{\otimes n}}^\lambda$  is the projection operator onto the isotypic indexed by  $\lambda$  appearing in  $(\mathbb{C}^d)^{\otimes n}$ . The sequence  $(E_n^{\text{spec}})_{n \in \mathbb{N}}$  is a spectral estimation scheme (see Fig. 4.1). In fact, this spectral estimation scheme can be derived from the state estimation scheme from Example 4.4.4 by pushing forward through the spectrum map  $\text{spec} : \mathcal{S}(\mathbb{C}^d) \rightarrow \Delta_d^\downarrow$  defined above. The corresponding rate function,  $I_\rho : \Delta_d^\downarrow \rightarrow [0, \infty]$ , for a fixed  $\rho \in \mathcal{S}(\mathbb{C}^d)$ , is for all spectra  $s \in \Delta_d^\downarrow$  equal to the relative entropy between  $s$  and the spectrum of  $\rho$  [KW01; Key06].

**Example 4.4.26.** Arguably the simplest example of an estimation scheme is the following estimation scheme when corresponds to the binary property,

$$\text{is}_\psi : \mathcal{S}(\mathcal{H}) \rightarrow \{F, T\}, \quad (4.4.71)$$

which decides whether the state  $\rho \in \mathcal{S}(\mathcal{H})$  is equal to a particular fixed, pure quantum state  $P_\psi \in \mathcal{S}(\mathcal{H})$  associated to the ray  $\psi \in \mathbb{P}\mathcal{H}$ , i.e.,

$$\text{is}_\psi(\rho) = \begin{cases} T & \rho = P_\psi, \\ F & \rho \neq P_\psi. \end{cases} \quad (4.4.72)$$

One method for determining the value of  $\text{is}_\psi(\rho)$  is to first perform the binary projective measurement  $\{P_\psi, P_{\psi^\perp}\}$  (where  $P_{\psi^\perp} := \mathbb{1}_{\mathcal{H}} - P_\psi$ ) independently on  $n \in \mathbb{N}$  identical copies of

the  $\rho$ , and then afterwards post-select on whether or not all of the outcomes obtained were  $P_\psi$ . The resulting  $(\text{is}_\psi)$ -estimation scheme, denoted by

$$E_n^\psi : \Sigma(\{F, T\}) \rightarrow \mathcal{B}(\mathcal{H}^{\otimes n}), \quad (4.4.73)$$

has the form [Hol11]

$$\begin{aligned} E_n^\psi(\{T\}) &= P_\psi^{\otimes n}, \\ E_n^\psi(\{F\}) &= \mathbb{1}_{\mathcal{H}}^{\otimes n} - P_\psi^{\otimes n}. \end{aligned} \quad (4.4.74)$$

For each state  $\rho$ , we have

$$\text{Tr}(E_n^\psi(\{T\})\rho^{\otimes n}) = \text{Tr}(P_\psi\rho)^n = \exp(-nI_\rho(T)), \quad (4.4.75)$$

and therefore the above  $(\text{is}_\psi)$ -estimation scheme satisfies the large deviation principle with rate function given by

$$\begin{aligned} I_\rho(T) &= -\log \text{Tr}(P_\psi\rho), \\ I_\rho(F) &= \begin{cases} \infty & \rho \neq P_\psi, \\ 0 & \rho = P_\psi. \end{cases} \end{aligned} \quad (4.4.76)$$

# Chapter 5

## Realizability problems

### 5.1 Realizability

The purpose of this section is to formally define the notion of a realizability problem for a property (or collection of properties) of a class of quantum states. Recall that a property of quantum states is considered to be a (measurable) function,  $f : \mathcal{S} \rightarrow X$ , mapping each quantum state,  $\rho \in \mathcal{S}$ , to their corresponding property value  $x = f(\rho) \in X$ . The problem of determining or characterizing which values the function  $f$  can take when evaluated over the set of all quantum states in  $\mathcal{S}$  is called the *realizability problem for  $f$* .

**Definition 5.1.1.** Given a measurable function  $f : \mathcal{S} \rightarrow X$  with  $(\mathcal{S}, \Sigma(\mathcal{S}))$  and  $(X, \Sigma(X))$  standard Borel spaces, define

- (i) the *inverse image* of  $f$  as  $f^{-1} : \Sigma(X) \rightarrow \Sigma(\mathcal{S})$  for  $\Delta \in \Sigma(X)$  by

$$f^{-1}(\Delta) = \{\rho \in \mathcal{S} \mid f(\rho) \in \Delta\}, \quad (5.1.1)$$

- (ii) and the *direct image* of  $f$ , by an abuse of notation, as  $f : \Sigma(\mathcal{S}) \rightarrow \Sigma(X)$  for  $S \in \Sigma(\mathcal{S})$  by

$$f(S) = \{x \in X \mid f^{-1}(x) \cap S \neq \emptyset\}. \quad (5.1.2)$$

The direct image,  $f$ , and inverse image,  $f^{-1}$ , of  $f$  form an adjoint pair in the sense that for all  $S \in \Sigma(\mathcal{S})$  and  $\Delta \in \Sigma(X)$ ,

$$f(S) \subseteq \Delta \iff S \subseteq f^{-1}(\Delta). \quad (5.1.3)$$

**Definition 5.1.2.** Given a function  $f : \mathcal{S} \rightarrow \mathcal{X}$ , the inverse image evaluated at a singleton  $\{x\} \in \Sigma(X)$ ,

$$f^{-1}(\{x\}) = \{\rho \in \mathcal{S} \mid f(\rho) = x\} \subseteq \mathcal{S}, \quad (5.1.4)$$

is called the *fiber over  $x$* .

**Remark 5.1.3.** Whenever the function  $f$  under consideration has codomain belonging to the set of real numbers,  $X \subseteq \mathbb{R}$ , the fiber over  $x \in \mathbb{R}$  may also be referred to as the *level-set* at  $x$ .

**Definition 5.1.4.** Given a property  $f : \mathcal{S} \rightarrow X$  and an element  $x \in X$ , the following conditions are equivalent:

- (i) there exists a  $\rho \in \mathcal{S}$  such that  $f(\rho) = x$ ,
- (ii) the fiber over  $x$  is non-empty, i.e.  $f^{-1}(\{x\}) \neq \emptyset$ ,
- (iii)  $x$  is an element of the direct image  $f(\mathcal{S})$  of  $\mathcal{S}$ .

If any, and thus all, of the above conditions are satisfied, the element  $x$  is said to be **realizable**. Otherwise,  $x$  is said to be **unrealizable**. The subset  $f(\mathcal{S}) \subseteq X$  of realizable elements will be referred to as the **realizable region** of  $f$ . The **realizability problem** for  $f$  is to decide, given  $x \in \mathcal{X}$ , whether  $x$  is realizable or unrealizable.

**Remark 5.1.5.** Since the fibers of any function are necessarily disjoint,

$$x \neq y \implies f^{-1}(\{x\}) \cap f^{-1}(\{y\}) = \emptyset, \quad (5.1.5)$$

the collection of non-empty fibers forms a *partition* of  $\mathcal{S}$  indexed by elements of the realizable region, denoted

$$\mathcal{S}/f := \{f^{-1}(\{x\}) \mid x \in f(\mathcal{S})\}. \quad (5.1.6)$$

Altogether, we have the following commutative diagram of maps

$$\begin{array}{ccc} \mathcal{S} & \xrightarrow{f} & \mathcal{X} \\ m_f \downarrow & & \uparrow \subseteq \\ \mathcal{S}/f & \xrightarrow{\cong} & f(\mathcal{S}) \end{array} \quad (5.1.7)$$

where the map  $m_f : \mathcal{S} \rightarrow \mathcal{S}/f$  sends each  $\rho \in \mathcal{S}$  to the unique non-empty fiber it belongs to:

$$m_f(\rho) = f^{-1}(\{f(\rho)\}). \quad (5.1.8)$$

**Example 5.1.6.** Consider  $\mathcal{S}$  to be set of density operators acting on a two-dimensional complex Hilbert space  $\mathcal{H} \cong \mathbb{C}^2$ , which may be identified with a point in the Bloch ball, e.g., by a point  $(x, y, z) \in \mathbb{R}^3$  such that  $\sqrt{x^2 + y^2 + z^2} \leq 1$  via the well-known bijection

$$\rho \mapsto (\text{tr}(\sigma_X \rho), \text{tr}(\sigma_Y \rho), \text{tr}(\sigma_Z \rho)) \quad (x, y, z) \mapsto \frac{1}{2}(I + x\sigma_X + y\sigma_Y + z\sigma_Z), \quad (5.1.9)$$

where  $\sigma_X, \sigma_Y$  and  $\sigma_Z$  are the usual Pauli-operators. A familiar example of a function of density operators that one might consider is the  $\sigma_Z$ -expectation value of  $\rho$ , i.e.  $\langle \sigma_Z \rangle(\rho) = \text{tr}(\sigma_Z \rho)$ . Evidently, the set of values for  $\text{tr}(\sigma_Z \rho)$  that can be realized by density operators is equal to the closed interval  $[-1, +1]$ , while the corresponding fibers over  $z \in [-1, +1]$ , as depicted by Fig. 5.1 (b), are closed disks of radius  $\sqrt{1 - z^2}$  in the  $xy$ -plane displaced by a height  $z$ . For a second example, consider the function  $\rho \mapsto \text{tr}(\rho^2)$ , otherwise known as the *purity* of  $\rho$ . Evaluated in  $(x, y, z)$ -coordinates, purity is equal to  $(1 + x^2 + y^2 + z^2)/2$  and thus the realizable region for purity is the closed interval  $[1/2, 1]$ . The fibers of purity, parameterized by  $p = \text{tr}(\rho^2)$ , correspond to spheres of radius  $\sqrt{2p - 1}$  as depicted by Fig. 5.1 (a).

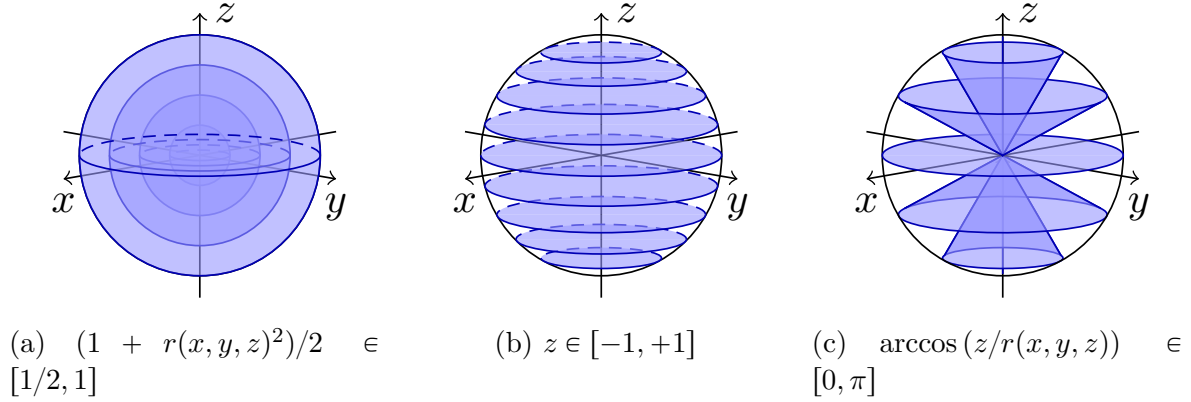


Figure 5.1: Various functions of the coordinates  $(x, y, z) \in \mathbb{R}^3$ , together with a depiction of their non-empty fibers in the unit ball. Also note  $r(x, y, z) = \sqrt{x^2 + y^2 + z^2}$ . In subfigure (c), the origin is treated as its own fiber and must be given its own special label distinct from those in  $[0, \pi]$ , such as  $\infty$ .

Before continuing, it will be important to elucidate the subtle, yet ultimately critical, distinction between *properties* and *parameters*. This claimed distinction is subtle simply because properties and parameters are often colloquially used to describe aspects or features of a system and in most settings, can be used interchangeably. Ultimately, in the setting investigated by this thesis, these two concepts can be seen to be dual to each other and thus cannot be used interchangeably. If the state or configuration of a system is described by some element of a space of possible states,  $\mathcal{S}$ , then a *parameterization* of the system is described by a function  $g : \Theta \rightarrow \mathcal{S}$  where the set  $\Theta$  (typically taken to be  $\mathbb{R}^d$  for some integer  $d$ ) is called the *parameter space* and its elements or coordinates called *parameters*. By comparison, the notion of a property emerges from considering a function  $f : \mathcal{S} \rightarrow X$  where the set  $X$  is called the *property space*. Any meaningful distinction between a property and a parameter, of course, disappears when the parameterization  $g : \Theta \rightarrow \mathcal{S}$  happens to be invertible, in which case its inverse,  $g^{-1} : \mathcal{S} \rightarrow \Theta$ , can be considered a property. In precisely the same manner, if a property  $f : \mathcal{S} \rightarrow X$  happens to be invertible, then its inverse,  $f^{-1} : X \rightarrow \mathcal{S}$ , can be considered as a parameterization. The focus of this thesis will be on properties which encode or represent some partial or incomplete information about the system and thus are, by assumption, *not* invertible. From this point of view, properties are considered to be conceptually distinct from parameters on the basis that to every parameter,  $\theta \in \Theta$ , there exists the state  $g(\theta) \in \mathcal{S}$ , whereas there might exist property values,  $x \in X$ , for which there does not exist a state possessing that property. In other words, unlike parameters, not all property values are *realizable*, and thus the problem of determining which property values are realizable becomes a non-trivial problem.

### 5.1.1 Quantitative realizability

Here we consider a variant of a realizability problem for the property  $f : \mathcal{S} \rightarrow X$  which aims to essentially quantify the proportion of states  $\rho \in \mathcal{S}$  that have the property value  $f(\rho) = x$  where  $x \in \Delta$  belongs to some measurable subset  $\Delta \subseteq X$ . The answer to this



question, of course, depends on a choice of normalized measure,  $\mu : \Sigma(\mathcal{S}) \rightarrow [0, 1]$ , over the set of states  $\rho \in \mathcal{S}$ , of which there are numerous physically motivated options [Woo90; ZS01]. Throughout this section, we assume  $\mathcal{S} = \mathcal{S}(\mathcal{H})$  is the set of density operators on a complex finite-dimensional Hilbert space  $\mathcal{H}$  equipped with topology induced by the operator norm on  $\text{End}(\mathcal{H})$  and let the  $\sigma$ -algebra over  $\mathcal{S}$  generated by that topology be denoted by  $\Sigma(\mathcal{S})$ .

**Remark 5.1.7.** Let  $f : \mathcal{S} \rightarrow X$  be a property of quantum states and let  $\mu : \Sigma(\mathcal{S}) \rightarrow [0, 1]$  be a prior probability measure over the set of states  $\mathcal{S}$ . The pushforward measure of  $\mu$  through the property  $f$ , is defined for  $\Delta \in \Sigma(X)$  by

$$(f_*\mu)(\Delta) = (\mu \circ f^{-1})(\Delta) = \mu(f^{-1}(\Delta)). \quad (5.1.10)$$

**Example 5.1.8.** For example, let  $\mathcal{H} = \mathbb{C}^d$  and let  $A \in \text{End}(\mathcal{H})$  be a Hermitian operator with non-degenerate spectra  $\{\lambda_1, \dots, \lambda_d\} \subset \mathbb{R}$ , and let the property under consideration be the  $A$ -expectation value,  $e_A : \mathcal{S} \rightarrow \mathbb{R}$  sending  $\rho \in \mathcal{S}$  to  $e_A(\rho) \in \mathbb{R}$ . Further suppose that the prior measure,  $\mu$ , only has support over the set of pure states,  $\mathbb{P}\mathcal{H}$ , and when  $\mu$  is restricted to its support, is equal to the normalized,  $U(d)$ -invariant Haar measure over  $\mathbb{P}\mathcal{H}$ . Then the pushforward of  $\mu$  through the  $A$  expectation value map,  $e_A$ , has density equal to (Theorem A.3.1 [Zha+22])

$$d(\mu \circ e_A^{-1})(a) = (d-1) \sum_{i=1}^d \frac{(a - \lambda_i)^{d-2} H(a - \lambda_i)}{\prod_{j \neq i} (\lambda_j - \lambda_i)} da. \quad (5.1.11)$$

where  $H(a)$  is the Heaviside function.

**Remark 5.1.9.** It is important to note that the support of the pushforward measure  $f_*\mu : \Sigma(X) \rightarrow [0, 1]$  is not necessarily equal to the image of the support of  $\mu : \Sigma(\mathcal{S}) \rightarrow [0, 1]$  through  $f : \mathcal{S} \rightarrow X$ , i.e.,  $f(\text{supp}(\mu)) \neq \text{supp}(f_*\mu)$ . For a simple example of this phenomena, consider the subset  $\mathcal{P} \subset \mathcal{S}$  of pure states and let  $f : \mathcal{S} \rightarrow \{T, F\}$  be the Boolean-valued property of whether a state is pure:

$$f(\rho) = \begin{cases} T & \rho \in \mathcal{P}, \\ F & \rho \notin \mathcal{P}. \end{cases} \quad (5.1.12)$$

Now consider a measure  $\mu : \Sigma(\mathcal{S}) \rightarrow [0, 1]$  for which the pure states are a null-set, i.e.  $\mu(\mathcal{P}) = 0$ , but nevertheless  $\mu$  has full-support on  $\mathcal{S}$  such as the Hilbert-Schmidt measure. The pushforward measure in this case satisfies

$$(f_*\mu)(T) = \mu(f^{-1}(T)) = \mu(\mathcal{P}) = 0, \quad (5.1.13)$$

$$(f_*\mu)(F) = \mu(f^{-1}(F)) = \mu(\mathcal{S} \setminus \mathcal{P}) = 1, \quad (5.1.14)$$

and thus its support is equal to  $\text{supp}(f_*\mu) = \{F\}$ . By comparison, since the support of  $\mu$  contains both pure and mixed states, we have  $f(\text{supp}(\mu)) = \{T, F\}$ . Consequently, the pushforward measure  $f_*\mu$  does not always faithfully capture the realizable region in general.

While a direct computation of the pushforward measure  $\mu \circ f^{-1}$  is certainly desirable, for a generic property  $f$ , this can be prohibitively difficult to do (see Remark A.3.3). Fortunately,

when provided with an  $f$ -estimation scheme,  $(E_n : \Sigma(X) \rightarrow \mathcal{B}(X^{\otimes n}))_{n \in \mathbb{N}}$ , it becomes possible to approximate  $\mu \circ f^{-1}$ . Recall that applying the POVM  $E_n$  to the state  $\rho^{\otimes n}$  yields a probability distribution, defined for  $\Delta \in \Sigma(X)$  by

$$\xi_n^\rho(\Delta) := \text{Tr}(E_n(\Delta)\rho^{\otimes n}). \quad (5.1.15)$$

When  $\mu$  represents some prior measure describing the identity of  $\rho \in \mathcal{S}$ , we encounter the notion of a de Finetti state.

**Definition 5.1.10.** Let  $\mu : \Sigma(\mathcal{S}) \rightarrow [0, 1]$  be a probability measure over  $\mathcal{S}$  and  $n \in \mathbb{N}$  and positive integer. The  **$\mu$ -de Finetti state** of degree  $n$  is

$$D_n^\mu := \int_{\mathcal{S}} \mu(d\rho)\rho^{\otimes n}. \quad (5.1.16)$$

**Proposition 5.1.11.** Let  $f : \mathcal{S} \rightarrow X$  be a property of quantum states, let  $(E_n : \Sigma(X) \rightarrow \mathcal{B}(\mathcal{H}^{\otimes n}))_{n \in \mathbb{N}}$  be an  $f$ -estimation scheme, and let  $\mu : \Sigma(\mathcal{S}) \rightarrow [0, 1]$  be a probability measure over the set of states  $\mathcal{S}$ . Then the sequence of probability measures  $(\Xi_n^\mu : \Sigma(X) \rightarrow [0, 1])_{n \in \mathbb{N}}$ , defined for all  $\Delta \in \Sigma(X)$  and  $n \in \mathbb{N}$  by

$$\Xi_n^\mu(\Delta) = \int_{\rho \in \mathcal{S}} \mu(d\rho)\text{Tr}(E_n(\Delta)\rho^{\otimes n}) = \text{Tr}(E_n(\Delta)D_n^\mu), \quad (5.1.17)$$

converges weakly to the pushforward measure of  $\mu$  through  $f$ . In other words, for all bounded continuous functions  $g : X \rightarrow \mathbb{R}$ ,

$$\int_{x \in X} g(x)\text{Tr}(dE_n(x)D_n^\mu) = \int_{\rho \in \mathcal{S}} g(f(\rho)) d\mu(\rho). \quad (5.1.18)$$

*Proof.* The proof follows from [Lemma 2.1.12](#) which states that the pushforward operation,  $f_*$ , is continuous with respect to weak topologies and thus respects the notion of weak convergence.  $\square$

**Remark 5.1.12.** If the sequence of measures,  $(\xi_n^\rho : \Sigma(X) \rightarrow [0, 1])_{n \in \mathbb{N}}$ , defined by [Eq. \(5.1.15\)](#) satisfies the large deviations principle with rate function  $I_\rho : X \rightarrow [0, 1]$ , then one might expect the sequence of measures,  $(\Xi_n^\mu : \Sigma(X) \rightarrow [0, 1])_{n \in \mathbb{N}}$ , defined by [Proposition 5.1.11](#), to satisfy the large deviations principle with rate function  $I' : X \rightarrow [0, 1]$ , defined for  $x \in X$  by

$$I'(x) := \inf_{\rho \in S_\mu} I_\rho(x) \quad (5.1.19)$$

where  $S_\mu = \text{supp}(\mu)$  is the support of the prior measure  $\mu$ . Unfortunately, to prove  $I'$  is indeed a rate function requires certain continuity assumptions of the function  $\rho \mapsto I_\rho(x)$  which may not hold in general. Moreover, to prove that  $(\Xi_n^\mu)_{n \in \mathbb{N}}$  satisfies the large deviation principle requires a more careful analysis of the uniformity with which large deviations principles for  $(\xi_n^\rho)_{n \in \mathbb{N}}$  hold with respect to  $\rho$ . See for example, Finnoff's integration theorem of sequences of probability kernels satisfying large deviation principles [[Fin02](#)].

### 5.1.2 Inclusivity of de Finetti states

**Definition 5.1.13.** Let  $\mu : \Sigma(\mathcal{S}) \rightarrow [0, 1]$  be a probability measure with support  $\text{supp}(\mu) \subseteq \mathcal{S}$  and define the function  $\gamma : \mathbb{N} \rightarrow \mathbb{R}_{\geq 0}$  for each  $n \in \mathbb{N}$  as

$$\gamma(n) := \inf\{c \in \mathbb{R}_{\geq 0} \mid \forall \sigma \in \text{supp}(\mu), cD_n^\mu \geq \sigma^{\otimes n}\}.$$

If  $\gamma$  does not increase too quickly, i.e.,

$$\lim_{n \rightarrow \infty} \gamma^{\frac{1}{n}}(n) = 1,$$

then  $\mu$  is said to be *inclusive* and the function  $\gamma$  is terms the *inclusivity function*.

**Definition 5.1.14.** Let  $n \in \mathbb{N}$  and let  $\mathcal{H}$  be a complex  $d$ -dimensional Hilbert space. Let projection operator onto the  $S_n$ -symmetric subspace  $\text{Sym}^n(\mathcal{H}) \subset \mathcal{H}^{\otimes n}$  be denoted by  $\Sigma_n$ ,

$$\Sigma_n = \frac{1}{n!} \sum_{\pi \in S_n} T_n(\pi). \quad (5.1.20)$$

The *normalized* state over the symmetric subspace will be expressed by

$$\underline{\Sigma}_n = \frac{\Sigma_n}{\text{Tr}(\Sigma_n)}, \quad (5.1.21)$$

The uniform symmetric state defined above constitutes the motivating example of an inclusive probability measure where the prior measure  $\mu$  is taken to be the  $U(d)$ -invariant Fubini-Study or Haar measure on  $\mathcal{S}$  with support over the pure states, i.e.,  $\text{supp}(\mu) = \{\rho \in \mathcal{S} \mid \text{rank}(\rho) = 1\}$ . In this case, the corresponding inclusivity function is given by the dimension of the  $n$ th symmetric subspace, i.e.,  $\gamma(n) = \binom{n+d-1}{n}$  where  $d = \dim(\mathcal{H})$ . Of course, this example can also be used to directly prove the inclusivity of the Hilbert-Schmidt measure with support over the entire set of density operators  $\mathcal{S}$  where the corresponding inclusivity measure is  $\gamma(n) = \binom{n+d^2-1}{n}$ . Moreover, it is conjectured that essentially all commonly encountered probability measures over the space of quantum states, such as those considered in [ZS01], are inclusive with respect to definition Definition 5.1.13. Perhaps the simplest examples of inclusive probability measures over quantum states are the discrete probability measures. For example, consider a discrete probability measure,  $\mu$ , with binary support  $\text{supp}(\mu) = \{\rho, \sigma\}$  of the form

$$\mu = \lambda\delta_\rho + (1 - \lambda)\delta_\sigma. \quad (5.1.22)$$

In this case, the inclusivity function,  $\gamma(n) = \max\{\frac{1}{\lambda}, \frac{1}{1-\lambda}\}$ , is independent of  $n$ .

**Theorem 5.1.15.** Let  $\mathcal{H}$  be a  $d$ -dimensional complex Hilbert space, let  $p \in \mathbb{N}$  be a positive integer and let

$$(X_n \in \mathcal{B}(\mathcal{H}^{\otimes pn}))_{n \in \mathbb{N}} \quad (5.1.23)$$

be a sequence of positive semidefinite operators. Let  $\underline{\Sigma}_{np}$  be the uniform state on the symmetric subspace  $\text{Sym}^{np}(\mathcal{H}) \subseteq \mathcal{H}^{\otimes np}$ . Further assume that the following limit exists for all  $\psi \in \mathbb{P}\mathcal{H}$ ,

$$C(\psi) := \limsup_{n \rightarrow \infty} \text{Tr}(X_n P_\psi^{\otimes pn})^{\frac{1}{n}}. \quad (5.1.24)$$

Then the supremum of  $C(\psi)$  over all  $\psi \in \mathbb{P}\mathcal{H}$  equals

$$\sup_{\psi \in \mathbb{P}\mathcal{H}} C(\psi) = \limsup_{n \rightarrow \infty} \text{Tr}(X_n \underline{\Sigma}_{pn})^{\frac{1}{n}} \quad (5.1.25)$$

$$(5.1.26)$$

*Proof.* Since the ray  $\psi^{\otimes pn}$  is a subspace of  $\text{Sym}^{pn}(\mathcal{H})$  and  $X_n$  is positive semidefinite, we have

$$\sup_{\psi \in \mathbb{P}\mathcal{H}} \text{Tr}(X_n P_\psi^{\otimes pn}) \leq \text{Tr}(X_n \underline{\Sigma}_{pn}). \quad (5.1.27)$$

Furthermore, since  $\text{Sym}^{pn}(\mathcal{H})$  supports an irreducible representation of  $\text{SU}(\mathcal{H})$ , we have by Schur's lemma

$$\text{Tr}(X_n \underline{\Sigma}_{pn}) = \dim(\text{Sym}^{pn}(\mathcal{H})) \int_{\phi \in \mathbb{P}\mathcal{H}} d\mu(\phi) \text{Tr}(X_n P_\phi^{\otimes pn}), \quad (5.1.28)$$

where  $\mu$  is the normalized Haar measure for  $\text{SU}(\mathcal{H})$ . Moreover, since  $\mu$  is a probability measure,

$$\int_{\phi \in \mathbb{P}\mathcal{H}} d\mu(\phi) \text{Tr}(X_n P_\phi^{\otimes pn}) \leq \sup_{\psi \in \mathbb{P}\mathcal{H}} \text{Tr}(X_n P_\psi^{\otimes pn}). \quad (5.1.29)$$

Finally, since  $\dim(\text{Sym}^{pn}(\mathcal{H})) = \binom{pn+d-1}{pn}$  grows at polynomial rate with increasing  $n$ ,

$$\lim_{n \rightarrow \infty} \dim(\text{Sym}^{pn}(\mathcal{H}))^{\frac{1}{pn}} = \lim_{n \rightarrow \infty} \dim(\text{Sym}^{pn}(\mathcal{H}))^{-\frac{1}{pn}} = 1. \quad (5.1.30)$$

Taking the supremum limit as  $n \rightarrow \infty$  then yields

$$\sup_{\psi \in \mathbb{P}\mathcal{H}} C(\psi) \leq \limsup_{n \rightarrow \infty} \text{Tr}(X_n \underline{\Sigma}_{pn})^{\frac{1}{n}} \leq \sup_{\psi \in \mathbb{P}\mathcal{H}} C(\psi), \quad (5.1.31)$$

which proves the claim.  $\square$

### 5.1.3 Moment polytopes

Consider a representation  $\Phi : G \rightarrow \text{GL}(\mathcal{H})$  with moment map  $\Omega_\Phi : \mathbb{P}\mathcal{H} \rightarrow i\mathfrak{k}^*$ . For which values of  $\omega = i\mathfrak{k}^*$  does there exist a ray  $\psi \in \mathbb{P}\mathcal{H}$  such that  $\Omega_\Phi(\psi) = \omega$ ?

To answer this question, our strategy is to make use of the correspondence, provided by [Proposition 4.3.11](#), between the  $\omega$ -capacity of a ray and the moment map of that ray equaling  $\omega$ . In this setting, for any given  $\psi$ , we have

$$\text{cap}_\Phi^\omega(\psi) = 1 \iff \Omega_\Phi(\psi) = \omega, \quad (5.1.32)$$

which, in turn, implies

$$\sup_{\psi \in \mathbb{P}\mathcal{H}} \text{cap}_\Phi^\omega(\psi) = 1 \iff \exists \psi \in \mathbb{P}\mathcal{H} : \Omega_\Phi(\psi) = \omega. \quad (5.1.33)$$

Therefore, the problem of determining which  $\omega$  belong to the image of the moment map  $\Omega_\Phi(\mathbb{P}\mathcal{H})$ , is equivalent to the problem of determining the maximum value of  $\text{cap}_\Phi^\omega(\psi)$ . Fortunately, the deformed strong duality result of [Theorem 4.3.14](#) offers an avenue for computing  $\text{cap}_\Phi^\omega(\psi)$  at least when  $\omega \in i\mathfrak{k}^*$  has a rational coadjoint orbit (where  $\omega = \text{Ad}^*(h)(\omega_+)$  and  $\ell\omega_+$  is dominant, analytically integral for some  $\ell$ ). Given an  $\omega \in i\mathfrak{k}^*$  of this form, [Theorem 4.3.14](#) yields,

$$\text{cap}_\Phi^\omega(\psi) = \limsup_{n \rightarrow \infty} \left\| Q_{\Phi^{\otimes n}}^\omega P_\psi^{\otimes n} \right\|_n^{\frac{1}{n}}, \quad (5.1.34)$$

where we have defined  $Q_{\Phi^{\otimes n}}^\omega$  as the projection operator onto the subspace of highest weight vectors for  $\Phi^{\otimes n}$  with weight  $n\omega_+$  rotated by  $h \in K$ , in accordance with [Definition 4.3.13](#),

$$Q_{\Phi^{\otimes n}}^\omega = \Phi^{\otimes n}(h) \Pi_{\Phi^{\otimes n}}^{n\omega_+} \Phi^{\otimes n}(h^{-1}). \quad (5.1.35)$$

In order to compute the supremum over all  $\psi \in \mathbb{P}\mathcal{H}$  we can make use of [Theorem 5.1.15](#) (for the special case where  $p = 1$  and  $X_n = Q_{\Phi^{\otimes n}}^\omega$ ) to obtain the following result.

**Proposition 5.1.16.** *Let  $\Omega_\Phi : \mathbb{P}\mathcal{H} \rightarrow i\mathfrak{k}^*$  be the moment map for a representation  $\Phi : G \rightarrow \text{GL}(\mathcal{H})$ , and let  $\omega \in i\mathfrak{k}^*$  have a rational coadjoint orbit. Then*

$$\limsup_{n \rightarrow \infty} \text{Tr}(Q_{\Phi^{\otimes n}}^\omega \underline{\Sigma}_n)^{\frac{1}{n}} = \sup_{\psi \in \mathbb{P}\mathcal{H}} \text{cap}_\Phi^\omega(\psi), \quad (5.1.36)$$

where  $\underline{\Sigma}_n$  is the uniform state on the  $S_n$ -symmetric subspace  $\text{Sym}^n(\mathcal{H}) \subseteq \mathcal{H}^{\otimes n}$ . Therefore, there exists  $\psi \in \mathbb{P}\mathcal{H}$  such that  $\Omega_\Phi(\psi) = \omega$  if and only if

$$\limsup_{n \rightarrow \infty} \text{Tr}(Q_{\Phi^{\otimes n}}^\omega \underline{\Sigma}_n)^{\frac{1}{n}} = 1. \quad (5.1.37)$$

In other words, the asymptotics of the quantity  $\text{Tr}(Q_{\Phi^{\otimes n}}^\omega \underline{\Sigma}_n)$  completely determines the realizability of  $\omega \in i\mathfrak{k}^*$  as the moment map of some state  $\psi \in \mathbb{P}\mathcal{H}$ . Using the symmetries of operator  $\underline{\Sigma}_n$ , one can simplify the result of [Proposition 5.1.16](#).

**Corollary 5.1.17.** *There exists  $\psi \in \mathbb{P}\mathcal{H}$  such that  $\Omega_\Phi(\psi) = \omega$  if and only if*

$$\limsup_{n \rightarrow \infty} \text{Tr}(\Pi_{\Phi^{\otimes n}}^{n\omega_+} \underline{\Sigma}_n)^{\frac{1}{n}} = 1, \quad (5.1.38)$$

where  $\Pi_{\Phi^{\otimes n}}^{n\omega_+}$  is the projection operator onto the subspace of highest weight vectors for  $\Phi^{\otimes n}$  with weight  $n\omega_+$ . In other words, the realizability of a rational  $\omega$  depends only on where its coadjoint intersects the fundamental Weyl chamber,  $\omega_+ \in C_+$ .

*Proof.* Since symmetric subspace supports an irreducible representation of  $U(\mathcal{H})$ ,  $\Phi$  restricted to the compact subgroup  $K$  is a unitary representation, we have for all  $k \in K$ ,

$$[\Phi^{\otimes n}(k), \underline{\Sigma}_n] = 0, \quad (5.1.39)$$

and therefore by [Eq. \(5.1.35\)](#) we conclude

$$\text{Tr}(Q_{\Phi^{\otimes n}}^\omega \underline{\Sigma}_n) = \text{Tr}(\Phi^{\otimes n}(h) \Pi_{\Phi^{\otimes n}}^{n\omega_+} \Phi^{\otimes n}(h^{-1}) \underline{\Sigma}_n) = \text{Tr}(\Pi_{\Phi^{\otimes n}}^{n\omega_+} \underline{\Sigma}_n). \quad (5.1.40)$$

The result then follows from [Eq. \(5.1.37\)](#) from [Corollary 5.1.17](#).  $\square$

Note that the statement of [Corollary 5.1.17](#) could have been anticipated in light of the fact that the moment map is  $K$ -equivariant ([Lemma 3.2.35](#)). Specifically, the realizability of  $\omega$  implies the realizability of  $\text{Ad}^*(k)(\omega)$  because

$$\Omega_\Phi(\psi) = \omega \implies \Omega_\Phi(k \cdot \psi) = \text{Ad}^*(k)(\omega). \quad (5.1.41)$$

An even more useful simplification arises from noticing the actions of  $S_n$  and  $\Phi^{\otimes n}$  on  $\mathcal{H}^{\otimes n}$  commute with each other. Therefore,  $[\Pi_{\Phi^{\otimes n}}^{n\omega_+}, \Sigma_n] = 0$ , and therefore, we have

$$\text{Tr}(\Pi_{\Phi^{\vee n}}^{n\omega_+}) = \text{Tr}(\Sigma_n \Pi_{\Phi^{\otimes n}}^{n\omega_+}), \quad (5.1.42)$$

where  $\Phi^{\vee n}$  is the  $n$ th symmetric power representation of  $G$  on  $\text{Sym}^n(\mathcal{H})$  from [Example 2.4.22](#) and  $\Pi_{\Phi^{\vee n}}^{n\omega_+}$  is the projection operator onto subspace of highest weight vectors with weight  $n\omega_+$  which are also  $S_n$ -symmetric (otherwise known as *covariants* of weight  $n\omega_+$  [[Wal+13](#)]). From this observation, one obtains a slightly more surprising result.

**Theorem 5.1.18.** *Let  $\Delta_\Phi$  denote the intersection of the image of the moment map with the positive Weyl chamber  $i\mathfrak{t}_+^*$ ,*

$$\Delta_\Phi := \{\omega_+ \in i\mathfrak{t}_+^* \mid \exists \psi : \Omega_\Phi(\psi) = \omega_+\}. \quad (5.1.43)$$

*Then  $\Delta_\Phi$  is a convex polytope, known as the **moment polytope** for  $\Phi$ .*

*Proof.* To prove  $\Delta_\Phi$  is a convex set, let  $\alpha_+, \beta_+ \in \Delta_\Phi$  be rational such that there exists  $\psi, \psi' \in \mathbb{P}\mathcal{H}$  satisfying

$$\Omega_\Phi(\psi) = \alpha_+, \quad \Omega_\Phi(\psi') = \beta_+. \quad (5.1.44)$$

Then for each  $a, b \in \mathbb{N}$  consider the projection operators  $\Pi_{\Phi^{\vee a}}^{a\alpha_+}$  and  $\Pi_{\Phi^{\vee b}}^{b\beta_+}$ . From [Corollary 5.1.17](#), [Eq. \(5.1.44\)](#) and [Eq. \(5.1.30\)](#), we conclude

$$\begin{aligned} \limsup_{a \rightarrow \infty} \text{Tr}(\Pi_{\Phi^{\vee a}}^{a\alpha_+})^{\frac{1}{a}} &= 1, \\ \limsup_{b \rightarrow \infty} \text{Tr}(\Pi_{\Phi^{\vee b}}^{b\beta_+})^{\frac{1}{b}} &= 1. \end{aligned} \quad (5.1.45)$$

More importantly, since  $\Pi_{\Phi^{\vee a}}^{a\alpha_+}$  and  $\Pi_{\Phi^{\vee b}}^{b\beta_+}$  are projection operators, their trace must be a non-negative integer, and therefore, [Eq. \(5.1.45\)](#) implies that there exists arbitrarily large  $a, b \in \mathbb{N}$  such that  $\text{Tr}(\Pi_{\Phi^{\vee a}}^{a\alpha_+}) \geq 1$  and  $\text{Tr}(\Pi_{\Phi^{\vee b}}^{b\beta_+}) \geq 1$ . Given a covariant of degree  $a$  with weight  $a\alpha_+$  and a covariant of degree  $b$  with weight  $b\beta_+$ , the symmetrization of their tensor product is necessarily a non-zero covariant of degree  $a + b$  with weight  $a\alpha_+ + b\beta_+$  and therefore, there exists arbitrarily large  $a, b \in \mathbb{N}$  such that

$$\text{Tr}(\Pi_{\Phi^{\vee(a+b)}}^{a\alpha_+ + b\beta_+}) \geq 1, \quad (5.1.46)$$

and similarly, for all  $n \in \mathbb{N}$ ,

$$\text{Tr}(\Pi_{\Phi^{\vee n(a+b)}}^{na\alpha_+ + nb\beta_+}) \geq 1. \quad (5.1.47)$$

Let  $\gamma_+$  be the convex combination of  $\alpha_+$  and  $\beta_+$ ,

$$\gamma_+ = \frac{a\alpha_+ + b\beta_+}{a + b}. \quad (5.1.48)$$

Then Eq. (5.1.47) implies

$$\limsup_{n \rightarrow \infty} \text{Tr}(\Pi_{\Phi^{\otimes n}}^{n\gamma_+})^{\frac{1}{n}} = 1, \quad (5.1.49)$$

which means, by Corollary 5.1.17, that  $\gamma_+ = \Omega_{\Phi}(\psi'')$  is the moment map of some  $\psi'' \in \mathbb{P}\mathcal{H}$  and thus  $\gamma_+ \in \Delta_{\Phi}$ . To further prove that  $\Delta_{\Phi}$  is a polytope requires a result which states the algebra of symmetric highest weight vectors, viewed as covariant polynomial functions in  $\mathcal{H}$ , is finitely generated [Wal14; Bri87].  $\square$

**Remark 5.1.19.** The moment polytope for  $\Phi$  provides a solution to the realizability problem for a given moment map. In particular,  $\omega \in i\mathfrak{k}^*$  is realizable as the moment map of some ray  $\psi \in \mathbb{P}\mathcal{H}$  if and only if the coadjoint orbit of  $\omega$ , denoted by  $\text{Ad}^*(K)(\omega) \subseteq i\mathfrak{k}^*$ , which intersects the positive Weyl chamber uniquely at  $\omega_+ \in i\mathfrak{t}_+^*$ , is such that  $\omega_+$  lies inside the moment polytope for  $\Phi$ :

$$\exists \psi \in \mathbb{P}\mathcal{H} : \Omega_{\Phi}(\psi) = \omega \iff \text{Ad}^*(K)(\omega) \cap \Delta_{\Phi} \neq \emptyset \iff \omega_+ \in \Delta_{\Phi}. \quad (5.1.50)$$

### 5.1.4 Beyond moment polytopes

The purpose of this section is to consider realizability problems which go beyond the realizability problem associated to a single moment map. Our approach is to consider the generalization of Proposition 5.1.16 to the case where one is interested in characterizing the image of a moment map  $\Omega_{\Phi_{(p)}} : \mathbb{P}(\mathcal{H}^{\otimes p}) \rightarrow i\mathfrak{k}^*$  with respect to a representations of the form  $\Phi_{(p)} : G \rightarrow \text{GL}(\mathcal{H}^{\otimes p})$ .

**Theorem 5.1.20.** *Let  $p \in \mathbb{N}$  be a positive integer and let  $\Phi_{(p)} : G \rightarrow \text{GL}(\mathcal{H}^{\otimes p})$  be a representation of a complex reductive group  $G$  on  $\mathcal{H}^{\otimes p}$ . Let  $\omega = \text{Ad}^*(h)(\omega_+) \in i\mathfrak{k}^*$  have rational coadjoint orbit and let  $Q_{\Phi^{\otimes n}}^{\omega}$  be the projection operator onto the subspace  $(\mathcal{H}^{\otimes p})^{\otimes n}$  defined by*

$$Q_{\Phi^{\otimes n}}^{\omega} = \Phi_{(p)}^{\otimes n}(h) \Pi_{\Phi_{(p)}^{\otimes n}}^{n\omega_+} \Phi_{(p)}^{\otimes n}(h^{-1}) \quad (5.1.51)$$

and where  $\Pi_{\Phi_{(p)}^{\otimes n}}^{n\omega_+}$  is the projection operator onto the highest weight subspace of  $(\mathcal{H}^{\otimes p})^{\otimes n}$  with weight  $n\omega_+$ . Then there exists a ray  $\psi \in \mathbb{P}\mathcal{H}$  such that

$$\Omega_{\Phi_{(p)}}(\psi^{\otimes p}) = \omega \in i\mathfrak{k}^*, \quad (5.1.52)$$

if and only if

$$\limsup_{n \rightarrow \infty} \text{Tr}(Q_{\Phi^{\otimes n}}^{\omega} \Sigma_{pm})^{\frac{1}{n}} = 1. \quad (5.1.53)$$

*Proof.* From Theorem 4.3.14, we have for all  $\Psi \in \mathbb{P}\mathcal{H}^{\otimes p}$ ,

$$\text{cap}_{\Phi_{(p)}}^{\omega}(\Psi) = \limsup_{n \rightarrow \infty} \text{Tr}(Q_{\Phi^{\otimes n}}^{\omega} \Psi^{\otimes n})^{\frac{1}{n}}. \quad (5.1.54)$$

Which implies for all  $\psi \in \mathbb{P}\mathcal{H}$ ,

$$\text{cap}_{\Phi_{(p)}}^{\omega}(\psi^{\otimes p}) = \limsup_{n \rightarrow \infty} \text{Tr}(Q_{\Phi^{\otimes n}}^{\omega} \psi^{\otimes p})^{\frac{1}{n}}. \quad (5.1.55)$$

Now applying [Theorem 5.1.15](#) to the case where  $X_n = Q_{\mathbb{F}^{\otimes n}}^\omega$  yields

$$\sup_{\psi \in \mathbb{P}\mathcal{H}} \text{cap}_{\Phi_{(p)}}^\omega(\psi^{\otimes p}) = \limsup_{n \rightarrow \infty} \text{Tr}(Q_{\mathbb{F}^{\otimes n}}^\omega \underline{\Sigma}_{pn})^{\frac{1}{n}}. \quad (5.1.56)$$

But  $\sup_{\psi \in \mathbb{P}\mathcal{H}} \text{cap}_{\Phi_{(p)}}^\omega(\psi^{\otimes p}) = 1$  if and only if there exists a  $\psi \in \mathbb{P}\mathcal{H}$  such that  $\Omega_{\Phi_{(p)}}(\psi^{\otimes p}) = \omega$ , which concludes the proof.  $\square$

**Remark 5.1.21.** In the special case where  $p = 1$ , [Theorem 5.1.20](#) reduces to [Proposition 5.1.16](#). Unlike [Proposition 5.1.16](#), however, the  $S_{np}$ -symmetric subspace is not necessarily invariant under the action of  $\Phi_{(p)}^{\otimes n}$  and therefore we generically have

$$[\Phi_{(p)}^{\otimes n}(k), \underline{\Sigma}_{np}] \neq 0. \quad (5.1.57)$$

Because of this lack of symmetry, we do not have a generalization of [Corollary 5.1.17](#) for  $p > 1$  and therefore no generalization of [Theorem 5.1.18](#).

Our main application of [Theorem 5.1.20](#) is to consider the case where  $\psi \in \mathbb{P}\mathcal{H}$

**Corollary 5.1.22.** *Consider two complex reductive groups,  $G_1$  and  $G_2$ , with representations acting on the same Hilbert space,*

$$\Phi_1 : G_1 \rightarrow \text{GL}(\mathcal{H}), \quad \Phi_2 : G_2 \rightarrow \text{GL}(\mathcal{H}), \quad (5.1.58)$$

*and assume  $\omega_1 \in i\mathfrak{k}_1^*$  and  $\omega_2 \in i\mathfrak{k}_2^*$  have rational coadjoint orbits. Then there exists a ray  $\psi \in \mathbb{P}\mathcal{H}$  satisfying*

$$\Omega_{\Phi_1}(\psi) = \omega_1, \quad \Omega_{\Phi_2}(\psi) = \omega_2, \quad (5.1.59)$$

*if and only if*

$$\limsup_{n \rightarrow \infty} \text{Tr}(\underline{\Sigma}_{2n}(Q_{\Phi_1^{\otimes n}}^{\omega_1} \otimes Q_{\Phi_2^{\otimes n}}^{\omega_2}))^{\frac{1}{2n}} = 1. \quad (5.1.60)$$

*Proof.* The proof follows from [Theorem 5.1.20](#) by setting  $p = 2$ ,  $G := G_1 \times G_2$  and  $\Phi_{(2)} = \Phi_1 \boxtimes \Phi_2$ , where

$$(\Phi_1 \boxtimes \Phi_2)(g_1, g_2) = \Phi_1(g_1) \boxtimes \Phi_2(g_2). \quad (5.1.61)$$

Then by [Lemma 3.2.38](#), we have

$$\Omega_{\Phi_1 \boxtimes \Phi_2}(\psi^{\otimes 2}) = \Omega_{\Phi_1}(\psi) \oplus \Omega_{\Phi_2}(\psi) \quad (5.1.62)$$

and moreover for any  $n \in \mathbb{N}$ ,

$$Q_{(\Phi_1 \boxtimes \Phi_2)^{\otimes n}}^{\omega_1 \oplus \omega_2} = Q_{\Phi_1^{\otimes n}}^{\omega_1} \otimes Q_{\Phi_2^{\otimes n}}^{\omega_2}. \quad (5.1.63)$$

$\square$

The proof technique from [Corollary 5.1.23](#) trivially generalizes to all  $p \in \mathbb{N}$ .

**Corollary 5.1.23.** *Let  $[p] = \{1, \dots, p\}$  be an index set and for each index  $j \in [p]$ , let  $\Phi_j : G_j \rightarrow \text{GL}(\mathcal{H})$  be a representation of a complex reductive group  $G_j$ , let  $\Omega_{\Phi_j} : \mathbb{P}\mathcal{H} \rightarrow (i\mathfrak{k}_j)^*$  be the moment map for  $\Phi_j$ , and let  $\omega_j \in i\mathfrak{k}_j^*$  have a rational coadjoint orbit. Then there exists a ray  $\psi \in \mathbb{P}\mathcal{H}$ , simultaneously satisfying*

$$\forall j \in [p] : \Omega_{\Phi_j}(\psi) = \omega_j, \quad (5.1.64)$$

*if and only if*

$$\limsup_{n \rightarrow \infty} \text{Tr}(\underline{\Sigma}_{np}(Q_{\Phi_1^{\otimes n}}^{\omega_1} \otimes \dots \otimes Q_{\Phi_p^{\otimes n}}^{\omega_p}))^{\frac{1}{np}} = 1. \quad (5.1.65)$$



### 5.1.5 The biriffle formula

In [Section 5.1.4](#), it was shown that the joint realizability of a given collection of properties of quantum states is entirely characterized by the asymptotics of quantities of the form

$$p_n(X_1, \dots, X_k) := \text{Tr}(\underline{\Sigma}_{nk}(X_1 \otimes X_2 \otimes \dots \otimes X_k)) \quad (5.1.66)$$

where  $(X_1, \dots, X_k)$  is a  $k$ -tuple of positive semidefinite operators in  $\text{End}(\mathcal{H}^{\otimes n})$  and  $\underline{\Sigma}_{nk}$  is the density operator on  $\text{End}(\mathcal{H}^{\otimes nk})$  associated to the maximally mixed state over the symmetric subspace  $\text{Sym}^{nk}(\mathcal{H}) \subseteq \mathcal{H}^{\otimes nk}$ . The purpose of this section is to describe a strategy for calculating the quantity appearing in [Eq. \(5.1.66\)](#) by relating the state  $\underline{\Sigma}_{nk}$  to a sum over the symmetric group  $S_{nk}$ , as outlined in [Section 5.1.2](#).

Our primary observation is to note that the density operator  $\underline{\Sigma}_{nk} \in \text{End}(\mathcal{H}^{\otimes nk})$ , which has the form

$$\underline{\Sigma}_{nk} = \frac{P_{\text{Sym}^{nk}(\mathcal{H})}}{\dim(\text{Sym}^{nk}(\mathcal{H}))}, \quad (5.1.67)$$

is fixed by the action  $T_{nk} : S_{nk} \rightarrow \text{GL}(\mathcal{H}^{\otimes nk})$  of the symmetric group  $S_{nk}$  on  $\mathcal{H}^{\otimes nk}$  meaning  $T_{nk}(\pi)\underline{\Sigma}_{nk} = \underline{\Sigma}_{nk} = \underline{\Sigma}_{nk}T_{nk}(\pi)$  for all  $\pi \in S_{nk}$ .

Therefore if one replaces each argument operator  $X_i \in \text{End}(\mathcal{H}^{\otimes n})$  with the  $S_n$ -fixed operator  $\tilde{X}_i := P_{\text{Sym}^n(\mathcal{H})}X_iP_{\text{Sym}^n(\mathcal{H})}$ , the value of  $p_n(X_1, \dots, X_k)$  is unchanged. Consequently, it will henceforth be assumed that each  $X_i$  in  $p_n(X_1, \dots, X_k)$  is fixed by  $S_n$  and therefore  $p_n(X_1, \dots, X_k)$  can be expressed as a sum over the double cosets  $S_n^{\times k} \backslash S_{nk} / S_n^{\times k}$  where  $S_n^{\times k}$  is viewed as a subset of  $S_{nk}$ :

$$S_n^{\times k} = S_n \times \dots \times S_n \subseteq S_{nk}. \quad (5.1.68)$$

While there are numerous injective group homomorphisms from  $S_n^{\times k}$  to  $S_{nk}$ , corresponding to the ways of partitioning the indices  $\{1, \dots, nk\}$  into  $k$  parts of equal size  $n$ , for the sake of concreteness, let the  $j$ -th component permutation  $\pi_j \in S_n$  in  $(\pi_1, \dots, \pi_k) \in S_n^{\times k}$  permute the  $j$ -th contiguous block of indices, i.e.  $\{(j-1)n+1, \dots, jn+1\}$ . The double cosets of  $S_n^{\times k} \backslash S_{nk} / S_n^{\times k}$  then correspond to the equivalence classes generated by the equivalence relation on  $S_{nk}$  defined by

$$\pi \sim \pi' \iff \exists \sigma, \sigma' \in S_n^{\times k} \subseteq S_{nk} : \pi = \sigma \circ \pi' \circ \sigma'. \quad (5.1.69)$$

It can be shown that the double cosets  $S_n^{\times k} \backslash S_{nk} / S_n^{\times k}$  are in bijection with the set of  $k \times k$  non-negative integer matrices,

$$\ell = \begin{pmatrix} \ell_{11} & \ell_{12} & \dots & \ell_{1k} \\ \ell_{21} & \ell_{22} & \dots & \ell_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ \ell_{k1} & \ell_{k2} & \dots & \ell_{kk} \end{pmatrix}, \quad (5.1.70)$$

whose row- and column- sums are equal to  $n$  (see [[Jon96](#), Theorem 2.2] or [[Ryb19](#)]) meaning:

$$\forall 1 \leq i, j \leq k : \sum_{i'=1}^k \ell_{i'j} = \sum_{j'=1}^k \ell_{ij'} = n. \quad (5.1.71)$$

Equivalently, the double cosets  $S_n^{\times k} \backslash S_{nk} / S_n^{\times k}$  are isomorphic to the set of directed multigraphs over  $k$  vertices such that the indegree and outdegree of each vertex is  $n$ .

Furthermore, for each double coset  $\ell \in S_n^{\times k} \backslash S_{nk} / S_n^{\times k}$ , there is a unique representative permutation belonging to  $\ell$ , which we denote by  $b_\ell \in S_{nk}$ , such that both  $b_\ell$  and its inverse  $b_\ell^{-1}$  preserve the order of indices within each block of size  $n$ , meaning

$$\begin{aligned} \forall j \in \{1, \dots, k\} : (j-1)n + 1 \leq i_1 < i_2 \leq jn + 1 \\ \implies (b_\ell(i_1) < b_\ell(i_2)) \cap (b_\ell^{-1}(i_1) < b_\ell^{-1}(i_2)). \end{aligned} \quad (5.1.72)$$

In other words, both  $b_\ell \in S_{nk}$  and its inverse,  $b_\ell^{-1}$  can be considered *riffle shuffles* of  $k$  packs of  $n$  cards. It is for this reason that we will refer to the forthcoming formula for  $p_n(X_1, \dots, X_k)$  as the *biriffle formula*.

The cardinality  $|\ell|$  of the double coset  $\ell \in S_n^{\times k} \backslash S_{nk} / S_n^{\times k}$  can also be calculated using the orbit-stabilizer theorem for double cosets:

$$|\ell| = \frac{|S_n^{\times k} \times S_n^{\times k}|}{|(S_n^{\times k} \times S_n^{\times k})_{b_\ell}|} = \frac{(n!)^{2k}}{\prod_{1 \leq i, j \leq k} \ell_{ij}!} \quad (5.1.73)$$

Next we apply the identity  $P_{\text{Sym}^{nk}(\mathcal{H})} = \frac{1}{(nk)!} \sum_{\pi \in S_{nk}} T_{nk}(\pi)$  to obtain

$$p_n(X_1, \dots, X_k) = \frac{1}{(nk)! \dim(\text{Sym}^{nk}(\mathcal{H}))} \sum_{\pi \in S_{nk}} \text{Tr}(T_{nk}(\pi)(X_1 \otimes \dots \otimes X_k)). \quad (5.1.74)$$

Since  $X_1 \otimes \dots \otimes X_k$  is fixed by  $S_n^{\times k}$ , the value of the summand for  $\pi \in S_{nk}$  equals the value of the summand for  $\pi' \in S_{nk}$  whenever  $\pi$  and  $\pi'$  belong to the same double coset  $\ell \in S_n^{\times k} \backslash S_{nk} / S_n^{\times k}$ . Altogether, we obtain the following result.

**Theorem 5.1.24.** *Let  $p_n(X_1, \dots, X_k)$  be defined as in Eq. (5.1.66) and assume that  $X_i \in \text{End}(\mathcal{H}^{\otimes n})$  is fixed by left and right action of  $S_n$  (via  $T_n : S_n \rightarrow \text{GL}(\mathcal{H}^{\otimes n})$ ) on  $X_i$  for all  $i \in [k]$ . Then the **biriffle formula** for  $p_n(X_1, \dots, X_k)$  is*

$$p_n(X_1, \dots, X_k) = \frac{(d-1)!}{(nk+d-1)!} \sum_{\ell \in S_n^{\times k} \backslash S_{nk} / S_n^{\times k}} |\ell| B_\ell(X_1, \dots, X_k) \quad (5.1.75)$$

where  $B_\ell(X_1, \dots, X_k)$ , called the **biriffle coupling**, is the scalar defined by

$$B_\ell(X_1, \dots, X_k) = \text{Tr}(T_{nk}(b_\ell)(X_1 \otimes \dots \otimes X_k)). \quad (5.1.76)$$

To better understand the nature of the biriffle coupling  $B_\ell(X_1, \dots, X_k)$ , note that  $b_\ell$ , and in particular  $T_{nk}(b_\ell)$ , can be interpreted, for entry  $\ell_{ij}$  in the matrix representation of  $\ell$  (Eq. (5.1.70)), as a partial contraction of  $\ell_{ij}$  tensor factors in the codomain of  $X_i$  with  $\ell_{ij}$  of the tensor factors in the domain of  $X_j$ . In other words,  $B_\ell(X_1, \dots, X_k)$  can be interpreted as a kind of trace over  $X_1 \otimes \dots \otimes X_k$  twisted by the coset  $\ell$ . Since each  $X_i$  is implicitly assumed to be fixed by the tensor-permutation action of  $S_n$  on  $\text{End}(\mathcal{H}^{\otimes n})$  this partial contraction can be performed over the symmetric subspace  $\text{Sym}^{\ell_{ij}}(\mathcal{H})$  of  $\mathcal{H}^{\otimes \ell_{ij}}$ .

To illustrate, consider that for any  $k$ -tuple of non-negative integers  $(q_1, \dots, q_k)$  with total  $Q = \sum_{j=1}^k q_j$ , we have a  $\text{GL}(\mathcal{H})$ -equivariant isometry representing the subspace inclusion relation between symmetric subspaces:

$$\iota_{q_1, \dots, q_k} : \text{Sym}^Q(\mathcal{H}) \hookrightarrow \bigotimes_{j=1}^k \text{Sym}^{q_j}(\mathcal{H}). \quad (5.1.77)$$

This isometry and its dual, denoted by  $\iota_{q_1, \dots, q_k}^*$ , can be represented using a diagrammatic shorthand (in the usual style of string diagrams, e.g. [Sel12; WBC11; Cvi08]).

$$\begin{array}{c} \begin{array}{c} \dots \\ \curvearrowright \\ q_1 \quad q_2 \quad \dots \quad q_k \\ \curvearrowleft \end{array} \\ \uparrow Q \\ \cong \iota_{q_1, \dots, q_k} : \text{Sym}^Q(\mathcal{H}) \hookrightarrow \bigotimes_{j=1}^k \text{Sym}^{q_j}(\mathcal{H}) \end{array} \quad (5.1.78)$$

$$\begin{array}{c} \uparrow Q \\ \cong \iota_{q_1, \dots, q_m}^* : \bigotimes_{j=1}^k \text{Sym}^{q_j}(\mathcal{H}) \twoheadrightarrow \text{Sym}^Q(\mathcal{H}). \\ \begin{array}{c} \curvearrowright \\ q_1 \quad q_2 \quad \dots \quad q_k \\ \curvearrowleft \end{array} \end{array} \quad (5.1.79)$$

Next let  $\tilde{X}_i$  denote the operator of the form

$$\tilde{X}_i : \bigotimes_{j=1}^k \text{Sym}^{\ell_{ji}}(\mathcal{H}) \rightarrow \bigotimes_{j'=1}^k \text{Sym}^{\ell_{ij'}}(\mathcal{H}) \quad (5.1.80)$$

defined by

$$\tilde{X}_i = \iota_{\ell_{i1}, \dots, \ell_{ik}}^* \circ X_i \circ \iota_{\ell_{1i}, \dots, \ell_{ki}}. \quad (5.1.81)$$

Then the coset coupling  $B_\ell(X_1, \dots, X_k)$  can be expressed diagrammatically, for small  $k$  as follows.

If  $k = 1$ , then  $\ell = (n)$  and thus

$$B_{(n)}(X_1) = \boxed{\tilde{X}_1} \Big|_n. \quad (5.1.82)$$

If  $k = 2$  then  $\ell = \begin{pmatrix} a & n-a \\ n-a & a \end{pmatrix}$  is parameterized by the integer  $a \in \{0, 1, \dots, n\}$  value and

$$B \begin{pmatrix} a & n-a \\ n-a & a \end{pmatrix} (X_1, X_2) = \begin{array}{c} \begin{array}{c} \circ \quad \circ \\ \text{---} \quad \text{---} \\ \circ \quad \circ \end{array} \\ \begin{array}{c} \downarrow k \quad \downarrow k \\ \boxed{\tilde{X}_2} \quad \boxed{\tilde{X}_1} \\ \uparrow k \quad \uparrow k \end{array} \\ \begin{array}{c} \circ \quad \circ \\ \text{---} \quad \text{---} \\ \circ \quad \circ \end{array} \end{array} \Big|_a. \quad (5.1.83)$$

If  $k = 3$  then,

$$B_\ell(X_1, X_2, X_3) = \text{Diagram} \quad (5.1.84)$$

In the special case where  $k = 2$ , the biriffle formula for  $p_n(X_1, X_2)$  leads to the following useful inequality on  $p_n(X_1, X_2)$ .

**Lemma 5.1.25.** *Let  $X_1, X_2$  be positive semidefinite operators on  $\mathcal{H}^{\otimes n}$  fixed by the tensor-permutation action  $T_n : S_n \rightarrow \text{GL}(\mathcal{H}^{\otimes n})$ . Then*

$$p_n(X_1, X_2) = \text{Tr}(\underline{\Sigma}_{2n}(X_1 \otimes X_2)) \geq \frac{\text{Tr}(\underline{\Sigma}_n(X_1 X_2))}{(2n)^{d-1}} \geq 0. \quad (5.1.85)$$

*Proof.* Using [Theorem 5.1.24](#) and [Eq. \(5.1.83\)](#), we have

$$p_n(X_1, X_2) = \text{Tr}(\underline{\Sigma}_{2n}(X_1 \otimes X_2)) = \binom{2n+d-1}{2n}^{-1} \frac{(n!)^2}{(2n)!} \sum_{a=0}^n \binom{n}{a}^2 \text{Diagram} \quad (5.1.86)$$

where the combinatoric coefficient arises from the decomposition  $S_{2n}$  into double cosets  $S_n^{\times 2} \backslash S_{2n} / S_n^{\times 2}$  and where [Eq. \(5.1.73\)](#) yields

$$\sum_{\ell \in S_n^{\times 2} \backslash S_{2n} / S_n^{\times 2}} |\ell| = \sum_{a=0}^n \frac{(n!)^4}{(a!)^2((n-a)!)^2} = (n!)^2 \sum_{a=0}^n \binom{n}{a}^2 = (2n)!. \quad (5.1.87)$$

By string diagrammatic manipulations, the factor involving  $X_1$  and  $X_2$  in the summand above can be written in terms of the transpose  $X_2^T$  of  $X_2$ :

$$\text{Diagram} = n \text{Diagram} = a \text{Diagram} \quad (5.1.88)$$

By assumption,  $X_1$  and  $X_2$  are positive semidefinite operators and thus  $X \otimes X_2^T$  is also positive semidefinite. Viewing the above formula as the trace of a positive semidefinite



Then, the moment map of the representation  $\Phi_S$  is a function of the form

$$\Omega_S : \mathbb{P}\mathcal{H} \rightarrow (\mathfrak{isu}(d))^*. \quad (5.2.3)$$

Using the moment map, we note a correspondence between  $\rho_S^\psi$  and  $\Omega_S(\psi)$ , captured by the following expression, which holds for all traceless Hermitian operators  $X \in \mathfrak{isu}(d)$ :

$$\mathrm{Tr}(\rho_S^\psi X) = \Omega_S(\psi)(X). \quad (5.2.4)$$

In other words, the moment map for the representation  $\Phi_S$  applied to  $\psi$  determines the reduced state of  $\psi$  on subsystem  $S$  up to normalization.

Now a pure state,  $\psi \in \mathbb{P}\mathcal{H}$ , is said to be *S-uniform* if its reduced state onto subsystem  $\mathcal{H}_S$  is proportional to the identity operator, meaning

$$\rho_S^\psi = \frac{\mathbb{1}_{\mathcal{H}_S}}{d_S}, \quad (5.2.5)$$

or equivalently, the moment map associated to the representation  $\Phi_S$  vanishes, i.e.,

$$\Omega_S(\psi) = 0. \quad (5.2.6)$$

A pure state,  $\psi \in \mathbb{P}\mathcal{H}$ , is said to be *absolutely maximally entangled*, if for all subsystems  $S \subseteq [p]$  of size

$$|S| \leq \left\lfloor \frac{D}{2} \right\rfloor, \quad (5.2.7)$$

the state  $\psi$  is *S-uniform*.

**Example 5.2.1.** For example, consider the case where  $p = 2$ ,  $d_1 = d_2 = 2$ , and  $D = 4$ . In this case, the familiar two-qubit Bell state  $\psi \in \mathbb{P}(\mathbb{C}^2 \otimes \mathbb{C}^2)$  associated to the unit vector

$$v := \frac{1}{\sqrt{2}}(e_0 \otimes e_0 + e_1 \otimes e_1) \in \mathbb{C}^2 \otimes \mathbb{C}^2, \quad (5.2.8)$$

is absolutely maximally entangled because both of its reduced states (on either the first or second subsystem) are proportional to identity matrix on  $\mathbb{C}^2$ , or equivalently, for all 2-dimensional traceless Hermitian operators  $X \in \mathfrak{isu}(2)$ ,

$$\begin{aligned} \mathrm{Tr}(\rho_1 X) &= \Omega_1(\psi)(X) = \langle v, (X \otimes \mathbb{1})v \rangle = 0, \\ \mathrm{Tr}(\rho_2 X) &= \Omega_2(\psi)(X) = \langle v, (\mathbb{1} \otimes X)v \rangle = 0. \end{aligned} \quad (5.2.9)$$

Using the strong-duality result from [Section 3.3.4](#), we see that  $\psi \in \mathbb{P}\mathcal{H}$  is *S-uniform*, i.e.  $\Omega_S(\psi) = 0$ , if and only if

$$\limsup_{n \rightarrow \infty} \mathrm{Tr}(P_\psi^{\otimes n} \Gamma_{\Phi_S^{\otimes n}})^{\frac{1}{n}} = 1, \quad (5.2.10)$$

where, for each  $n \in \mathbb{N}$ ,  $\Gamma_{\Phi_S^{\otimes n}}$  is the projection operator onto the subspace of vectors in  $(\mathcal{H}_S \otimes \mathcal{H}_{-S})^{\otimes n}$  which are invariant under the action of  $\mathrm{SU}(d_S)$  through the representation  $\Phi_S$ .

Combining this observation with the results from [Section 5.1.4](#), we obtain the following result which may be used to determine, for any given collection of subsystems  $S_1, \dots, S_m \subseteq [p]$ , whether or not there exists a state which is  $S_j$ -uniform for all  $j \in \{1, \dots, m\}$ .

**Corollary 5.2.2.** *Let  $S_1, \dots, S_m \subseteq [p]$  be a collection of subsystems indexed by  $j \in [m] \in \{1, \dots, m\}$ . Then, for each  $n \in \mathbb{N}$  and  $j \in [m]$ , let  $\Gamma_{\Phi_{S_j}^{\otimes n}}$  be the aforementioned projection operator onto the subspace of vectors in  $(\mathbb{C}^D)^{\otimes n}$  invariant under the action of  $\Phi_{S_j}^{\otimes n} : \text{SU}(d_j) \rightarrow \text{SU}(D^n)$  defined by*

$$\Phi_{S_j}(U_{S_j}) := U_{S_j}^{\otimes n} \otimes \mathbb{1}_{-S_j}^{\otimes n}. \quad (5.2.11)$$

*Then there exists a ray  $\psi \in \mathbb{P}(\mathbb{C}^D)$  which is  $S_j$ -uniform for all  $j \in [m]$  if and only if*

$$\limsup_{n \rightarrow \infty} \text{Tr}(\underline{\Sigma}_{nm}(\Gamma_{\Phi_{S_1}^{\otimes n}} \otimes \dots \otimes \Gamma_{\Phi_{S_m}^{\otimes n}})) = 1 \quad (5.2.12)$$

where  $\underline{\Sigma}_{nm}$  is defined in [Definition 5.1.14](#).

This result thus constitutes a first step toward, at least partially, generalizing the result of Bryan, Reichstein, and Van Raamsdonk [\[BRV18\]](#) for the complete solution to the existence or non-existence of *locally* maximally entangled states (i.e., the case where  $|S_j| = 1$  for all  $j$ ) as a function of local dimensions.

## 5.2.2 Quantum marginals

The quantum marginals problem can be understood as a non-uniform generalization of the problem considered in [Section 5.2.1](#). In fact, by using the marginal representations defined by [Eq. \(5.2.2\)](#), we observe that [Corollary 5.1.23](#) already provides an asymptotic solution to the quantum marginals problem, at least for marginal states with rational spectra. In [Chapter 6](#), we will revisit the quantum marginals problem in greater detail and will encounter an alternative, but ultimately related, strategy for asymptotically characterizing which collections of marginals are jointly realizable.

# Chapter 6

## Quantum marginal problems

The purpose of this chapter is to explore a class of realizability problems known as quantum marginal problems. From the perspective of [Chapter 5](#), and for the purposes of this chapter, a quantum marginal problem can be understood as a realizability problem where the properties under investigation are *localized* in the sense that they pertain to subsystems of a multipartite Hilbert space. Although a solution to the quantum marginal problem was already presented in [Section 5.2.2](#), this chapter provides an alternative, albeit related, solution in the form of [Theorem 6.3.4](#) which has the advantage of being much easier to articulate. The contents of this chapter are taken directly from my paper on the quantum marginals problem [[Fra22](#)], with only minor modifications.

### 6.1 Introduction

The quantum marginal problem (QMP) is interested in characterizing the space of reduced/marginal states of a multipartite quantum state, and is widely regarded as being an important, albeit challenging problem to solve.

Connections and applications of the QMP to other topics in quantum theory (and beyond) include multipartite entanglement and separability [[CKW00](#); [Wal+13](#)], quantum error correction [[Ock+11](#); [Hub17](#); [Yu+21](#)], entropic constraints [[Maj18](#); [CŞW18](#); [Kim20](#); [Osb08](#)], other forms of quantum realizability problems [[HMZ16](#); [Haa+21](#); [Doh+08](#)], the asymptotics of representations of the symmetric groups [[DH05](#); [CM06](#); [CŞW18](#)], the asymptotic restriction problem for tensors [[CVZ18](#)], random matrix theory [[CM21](#); [Chr+14](#)], and many-body physics [[Col01](#); [Sch15](#)].

The quantum marginal problem, and its associated terminology, is derived from an analogous problem in probability theory called the *classical* marginal problem. The classical marginal problem is concerned with characterizing the relationships between the various marginal distributions of a joint, multivariate distribution [[FC12](#); [Vor62](#); [Mal88](#)]. As a joint probability distribution and its marginals can always be faithfully represented by the eigenvalues of a joint quantum state and its marginals using a product eigenbasis, the quantum marginal problem subsumes the classical marginal problem, and consequently, any of its applications [[FC12](#); [LSW11](#); [AB11](#); [FW18](#)].

One of the earliest formulations of the problem dates back to the early 1960s, when,



for the purposes of simplifying calculations involving the atomic and molecular structure, quantum chemists became interested in characterizing the possible reduced density matrices of a system of  $N$  interacting fermions [Cou60; Col63]. This version of the problem, referred to as the  $N$ -representability problem, has a long history [CY00; Col01; LTC+13; BD72; Rus07; Kly09] that continues to evolve [Maz12b; Maz12a; Kly06; Cas+21].

Now the QMP comes in a variety of flavours which can be broadly organized by considering any additional assumptions or constraints that are imposed on either i) the properties of the joint state, and/or ii) the properties of the set of candidate density operators [TV15].

When focusing on the joint state, specializations of the QMP exist where the joint state is assumed to be fermionic [CY00; SGC13], bosonic [WMN10], Gaussian [Eis+08; Vla15], separable [NBA21], or having symmetric eigenvectors [AFT20]. Generally speaking, the QMP is difficult in the sense that it is a QMA-complete problem [Liu06; LCV07; WMN10; Boo12]. For the purposes of this section, the only restriction imposed on joint states will be that they live in a finite-dimensional Hilbert space, with a finite and fixed number of subsystems.

Regarding the list of candidate density operators, e.g.,  $(\rho_{AB}, \rho_{BC}, \rho_{AC})$ , the list of subsystems they correspond to, e.g.,  $(AB, BC, AC)$ , is known as the *marginal scenario*, while the individual elements, e.g.,  $AB$ ,  $BC$ , or  $AC$ , will be referred to as *marginal contexts*. A key consideration for understanding previous work on the QMP is whether the marginal contexts are disjoint. When the marginal contexts are disjoint, a complete solution to the QMP is known [Kly04; KS06]. For a given specification of Hilbert space dimensions, this solution takes the form of a finite list of linear inequality constraints on the spectra of the candidate density operators. These solutions furthermore recover earlier results pertaining to low dimensional Hilbert-spaces [HSS03; Hig03; Bra03; HZG05].

In contrast, when the marginal scenario involves overlapping marginal contexts, existing results are comparatively more sporadic and typically weaker, being only applicable to low-dimensional systems, small numbers of parties, or only yielding necessary but insufficient constraints [Che+14; CLL13; BSS06; Hal07; Che+16]. One promising approach, developed in [CŞW18] for relating marginal spectra to the recoupling theory of the symmetric group, appears limited to situations where the marginal contexts do not overlap too much.

Whenever a candidate set of density operators is explicitly given, one strategy to decide their realizability is to use convex optimization techniques, e.g., semidefinite programming [VB96]. If the joint state is not necessarily pure, realizability can be decided with a single semidefinite program [Hal07]. Additionally, when the joint state is assumed pure, realizability can still be decided by an infinite hierarchy of semidefinite programs [Yu+21]. In either case, analytic inequality constraints that serve as witnesses for unrealizability can be extracted from the outputs of such semidefinite programs [Hal07].

The objective of this chapter is to improve our understanding of the QMP, in particular for the case of overlapping marginal contexts, by i) deriving inequality constraints that are necessarily satisfied by all realizable collections of density operators, and ii) proving that if a collection of density operators satisfies these inequalities, then they are realizable. This chapter begins by formally defining the QMP and then reformulating it from a different perspective. The primary advantage of this reformulation of the QMP is that it exposes an implicit symmetry of the problem which is helpful in deriving our main result.

## 6.2 A Reformulation of the QMP

This section introduces some notation and terminology that is sufficient to formally define both the QMP and an equivalent reformulation that is better suited for the techniques developed in subsequent sections.

First and foremost, every Hilbert space considered will be complex, finite-dimensional, and labeled by some subscript  $X$ , e.g.,  $\mathcal{H}_X$ . For each labeled Hilbert space,  $\mathcal{H}_X$ , we will implicitly assume there exists some canonical orthonormal basis such that  $\mathcal{H}_X \cong \mathbb{C}^{d_X}$  where  $d_X = \dim(\mathcal{H}_X)$ . The corresponding set of linear operators, density operators (states), and pure states of  $\mathcal{H}_X$  are respectively denoted  $\mathcal{L}(\mathcal{H}_X)$ ,  $\mathcal{D}(\mathcal{H}_X)$ , and  $\mathbb{P}(\mathcal{H}_X)$ . The identity operator on  $\mathcal{H}_X$  is denoted  $\mathbb{1}_X$ .

Given a list of labels,  $S = (X_1, \dots, X_k)$ , which we identify with the concatenated string,  $S \simeq X_1 \cdots X_k$ , the composite Hilbert space  $\mathcal{H}_{X_1} \otimes \cdots \otimes \mathcal{H}_{X_k}$ , will be labeled by  $S$  itself and thus denoted  $\mathcal{H}_S$ . For instance, if  $S = ABC$ , then  $\mathcal{H}_{ABC} = \mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C$ . Additionally, for any positive integer  $n$  and label  $X$ , the list of labels consisting of  $n$  copies of  $X$  will be abbreviated by

$$nX := (X, \overset{n}{\cdot}, X) \simeq X \overset{n}{\cdot} X \quad (6.2.1)$$

Using this notational convention, the  $n$ th tensor-power of a Hilbert space  $\mathcal{H}_X$  can be written as  $\mathcal{H}_X^{\otimes n} = \mathcal{H}_{nX} = \mathcal{H}_{X \overset{n}{\cdot} X}$ .

Associated to any instance of the QMP, is a *joint* (or *global*) Hilbert space  $\mathcal{H}_J$  where  $J = (X_1, \dots, X_p)$  is a given finite list of labels called the *joint context*. Every non-empty sublist  $S$  of  $J$  will be called a *marginal context*. For each  $S \subseteq J$ , the partial trace from  $\mathcal{H}_J$  onto  $\mathcal{H}_S$  will be denoted by

$$\text{Tr}_{J \setminus S} : \mathcal{L}(\mathcal{H}_J) \rightarrow \mathcal{L}(\mathcal{H}_S). \quad (6.2.2)$$

A finite, non-empty tuple of marginal contexts,

$$\mathcal{M} = (S_1, \dots, S_m), \quad (6.2.3)$$

is called a *marginal scenario*. The cardinality of a marginal scenario will always be denoted by  $m = |\mathcal{M}|$ .

**Problem 6.2.1** (QMP). Given a marginal scenario,  $\mathcal{M} = (S_1, \dots, S_m)$ , and list of states  $(\rho_{S_1}, \dots, \rho_{S_m})$  where  $\rho_{S_i} \in \mathcal{D}(\mathcal{H}_{S_i})$  for each  $i \in \{1, \dots, m\}$ , decide if there exists a joint pure state  $\psi_J \in \mathbb{P}(\mathcal{H}_J)$  such that

$$\forall S_i \in \mathcal{M} : \rho_{S_i} = \text{Tr}_{J \setminus S_i}(\psi_J). \quad (6.2.4)$$

The assumption above of purity for the joint state is made without loss of generality as shown in [Section 6.6.1](#).

Whenever such a pure state  $\psi_J$  exists, the  $m$ -tuple of states  $(\rho_{S_1}, \dots, \rho_{S_m})$  is said to be *realizable* and otherwise they are *unrealizable*.

For the sake of brevity, unless otherwise specified, a joint Hilbert space  $\mathcal{H}_J$  with joint context  $J$  will always be implicitly given together with a particular marginal scenario  $\mathcal{M} = (S_1, \dots, S_m)$  of length  $m$ .

Our first step is to reinterpret the  $m$  linear constraints imposed on the joint state  $\psi_J$  by [Eq. \(6.2.4\)](#) as a single linear constraint on the  $m$ th tensor-power state,  $\psi_J^{\otimes m}$ . Specifically,

the  $m$ -tuple of states  $(\rho_{S_1}, \dots, \rho_{S_m})$  satisfies Eq. (6.2.4) if and only if

$$\begin{aligned} \rho_{S_1} \otimes \cdots \otimes \rho_{S_m} &= \text{Tr}_{J \setminus S_1}(\psi_J) \otimes \cdots \otimes \text{Tr}_{J \setminus S_m}(\psi_J), \\ &= (\text{Tr}_{J \setminus S_1} \otimes \cdots \otimes \text{Tr}_{J \setminus S_m})(\psi_J^{\otimes m}). \end{aligned} \quad (6.2.5)$$

This observation motivates the following definitions.

**Definition 6.2.2.** The *Hilbert space on  $\mathcal{M}$* , denoted  $\mathcal{H}_{\mathcal{M}}$ , is

$$\mathcal{H}_{\mathcal{M}} := \mathcal{H}_{S_1} \otimes \cdots \otimes \mathcal{H}_{S_m}. \quad (6.2.6)$$

The *partial trace from  $mJ$  onto  $\mathcal{M}$* , denoted  $\text{Tr}_{mJ \setminus \mathcal{M}}$ , is

$$\text{Tr}_{mJ \setminus \mathcal{M}} := \text{Tr}_{J \setminus S_1} \otimes \cdots \otimes \text{Tr}_{J \setminus S_m}. \quad (6.2.7)$$

Note that the partial trace from  $mJ$  onto  $\mathcal{M}$ ,  $\text{Tr}_{mJ \setminus \mathcal{M}}$ , is simply the partial trace operation mapping elements of  $\mathcal{L}(\mathcal{H}_{mJ}) = \mathcal{L}(\mathcal{H}_J^{\otimes m})$  to elements of  $\mathcal{L}(\mathcal{H}_{\mathcal{M}}) = \mathcal{L}(\mathcal{H}_{S_1} \otimes \cdots \otimes \mathcal{H}_{S_m})$ .

**Definition 6.2.3.** An  $\mathcal{M}$ -*product state* is any state,  $\rho_{\mathcal{M}} \in \mathcal{D}(\mathcal{H}_{\mathcal{M}})$ , of the form

$$\rho_{\mathcal{M}} = \rho_{S_1} \otimes \cdots \otimes \rho_{S_m} \quad (6.2.8)$$

where each component,  $\rho_{S_i}$ , is a state in  $\mathcal{D}(\mathcal{H}_{S_i})$ .

Since there is a bijection between  $m$ -tuples of states  $(\rho_{S_1}, \dots, \rho_{S_m})$  and  $\mathcal{M}$ -product states  $\rho_{S_1} \otimes \cdots \otimes \rho_{S_m}$ , the QMP can be equivalently restated entirely in terms of  $\rho_{\mathcal{M}}$ .

**Problem 6.2.4.** Given an  $\mathcal{M}$ -product state,  $\rho_{\mathcal{M}}$ , determine whether or not there exists a pure state  $\psi_J$  such that

$$\rho_{\mathcal{M}} = \text{Tr}_{mJ \setminus \mathcal{M}}(\psi_J^{\otimes m}). \quad (6.2.9)$$

The equivalence between Problem 6.2.1 and Problem 6.2.4 follows from Eq. (6.2.5); moreover, whenever such a pure state  $\psi_J$  exists in either formulation of the QMP, it satisfies both Eq. (6.2.4) and Eq. (6.2.9). Pursuant to this equivalence, an  $\mathcal{M}$ -product state,  $\rho_{S_1} \otimes \cdots \otimes \rho_{S_m}$ , is said to be *realizable* whenever the  $m$ -tuple of states  $(\rho_{S_1}, \dots, \rho_{S_m})$  is realizable (and *unrealizable* otherwise). The set of all realizable  $\mathcal{M}$ -product states will be denoted  $\mathcal{C}_{\mathcal{M}}$ .

## 6.3 Necessary and Sufficient Inequality Constraints

In this section, we construct inequalities that are necessarily satisfied by *all* realizable  $\mathcal{M}$ -product states,  $\rho_{\mathcal{M}}$ . These inequalities, therefore, can be used to answer the QMP in the negative; if an  $\mathcal{M}$ -product state violates any of the forthcoming inequalities, then it is unrealizable. In addition, it will be shown that if an  $\mathcal{M}$ -product state,  $\rho_{\mathcal{M}}$ , satisfies all of the forthcoming inequalities, then it must be realizable.

These inequalities emerge from considering the permutation symmetry of the  $k$ th tensor power,  $\psi_J^{\otimes k}$ , of a pure state  $\psi_J \in \mathbb{P}(\mathcal{H}_J)$ . For each  $k \in \mathbb{N}$ , let  $\Sigma_k$  be the symmetric group on  $k$  symbols, and let  $T_J : \Sigma_k \rightarrow \mathcal{L}(\mathcal{H}_J^{\otimes k})$  be the representation of  $\Sigma_k$  acting on  $\mathcal{H}_J^{\otimes k}$  by permutation of its  $k$  factors.

**Definition 6.3.1.** The  $k$ th symmetric subspace  $\vee^k \mathcal{H}_J \subseteq \mathcal{H}_J^{\otimes k}$  is defined as

$$\vee^k \mathcal{H}_J = \{|\phi\rangle \in \mathcal{H}_J^{\otimes k} \mid \forall \pi \in \Sigma_k, T_J(\pi)|\phi\rangle = |\phi\rangle\}. \quad (6.3.1)$$

The orthogonal projection operator onto  $\vee^k \mathcal{H}_J$  will be denoted by  $\Pi_J^{(k)}$ .

Given any vector  $|\psi_J\rangle \in \mathcal{H}_J$ , it is straightforward to verify that  $|\psi_J\rangle^{\otimes k}$  is an element of the  $k$ th symmetric subspace  $\vee^k \mathcal{H}_J \subseteq \mathcal{H}_J^{\otimes k}$ .

**Proposition 6.3.2.** Let  $\psi_J = |\psi_J\rangle\langle\psi_J|$  be a pure state, let  $k \in \mathbb{N}$ . Then

$$\psi_J^{\otimes k} \leq \Pi_J^{(k)}, \quad (6.3.2)$$

where  $\Pi_J^{(k)}$  is defined in [Definition 6.3.1](#).

Throughout this chapter, an inequality  $A \geq B$  between Hermitian operators  $A$  and  $B$  always indicates that  $A - B$  is positive semidefinite, i.e.  $A - B \geq 0$ . See [[Bha97](#), Section V].

By comparing [Eq. \(6.2.9\)](#) with [Eq. \(6.3.2\)](#), and recalling that partial traces are positive channels, it becomes possible to *eliminate*  $\psi_J$  from [Eq. \(6.2.9\)](#). For example, when  $k = m$ , the partial trace  $\text{Tr}_{mJ \setminus \mathcal{M}}$  applied to [Eq. \(6.3.2\)](#) yields  $\text{Tr}_{mJ \setminus \mathcal{M}}(\psi_J^{\otimes m}) \leq \text{Tr}_{mJ \setminus \mathcal{M}}(\Pi_J^{(m)})$  and thus we obtain the following corollary.

**Corollary 6.3.3.** If  $\rho_{\mathcal{M}}$  is a realizable  $\mathcal{M}$ -product state, then

$$\rho_{\mathcal{M}} \leq \text{Tr}_{mJ \setminus \mathcal{M}}(\Pi_J^{(m)}). \quad (6.3.3)$$

In [Section 6.6.3](#), it is shown that the utility of this constraint is quite sensitive to the marginal scenario under consideration. For certain marginal scenarios, [Eq. \(6.3.3\)](#) happens to be satisfied by all  $\mathcal{M}$ -product states, and thus is useless for the purposes of the QMP. Nevertheless, for other marginal scenarios, [Eq. \(6.3.3\)](#) happens to be violated by some  $\mathcal{M}$ -product states, and thus it serves as a non-trivial condition for the realizability of an  $\mathcal{M}$ -product state  $\rho_{\mathcal{M}}$ .

Analogous reasoning can be used to construct stronger inequality constraints for the QMP. When  $k$  is a multiple of  $m$ ,  $k = nm$ , one can apply the  $n$ th tensor power of  $\text{Tr}_{mJ \setminus \mathcal{M}}$  to both sides of [Eq. \(6.3.2\)](#). While it is clear from the preceding discussion that this will yield inequality constraints necessarily satisfied by all  $\mathcal{M}$ -product states, we will additionally show that their satisfaction for all  $n \in \mathbb{N}$  is *sufficient* to conclude that  $\rho_{\mathcal{M}}$  is a realizable  $\mathcal{M}$ -product state.

**Theorem 6.3.4.** An  $\mathcal{M}$ -product state,  $\rho_{\mathcal{M}}$ , is realizable if and only if for all  $n \in \mathbb{N}$ ,

$$\rho_{\mathcal{M}}^{\otimes n} \leq \text{Tr}_{mJ \setminus \mathcal{M}}^{\otimes n}(\Pi_J^{(nm)}). \quad (6.3.4)$$

Note that  $\text{Tr}_{mJ \setminus \mathcal{M}}^{\otimes n}$  is the partial trace operation taking elements of  $\mathcal{L}(\mathcal{H}_J^{\otimes nm})$  to elements of  $\mathcal{L}(\mathcal{H}_{\mathcal{M}}^{\otimes n})$ .

To prove [Theorem 6.3.4](#), we first need the following lemma which states that the upper-bound in [Eq. \(6.3.4\)](#) represents the expected value of  $\sigma_{\mathcal{M}}^{\otimes n}$ , up to normalization, when  $\sigma_{\mathcal{M}}$  is sampled according to a probability measure,  $\nu_{\mathcal{M}}$ , whose support is precisely the set of realizable  $\mathcal{M}$ -product states, denoted  $\mathcal{C}_{\mathcal{M}}$ .

**Lemma 6.3.5.** *There exists a probability measure,  $\nu_{\mathcal{M}}$ , over  $\mathcal{D}(\mathcal{H}_{\mathcal{M}})$ , with support  $\mathcal{C}_{\mathcal{M}}$ , such that for all  $n \in \mathbb{N}$ ,*

$$\mathrm{Tr}_{mJ \setminus \mathcal{M}}^{\otimes n}(\Pi_J^{(nm)}) = \binom{nm+d_J-1}{nm} \int_{\mathcal{C}_{\mathcal{M}}} \nu_{\mathcal{M}}(d\sigma_{\mathcal{M}}) \sigma_{\mathcal{M}}^{\otimes n}, \quad (6.3.5)$$

where  $d_J = \dim(\mathcal{H}_J)$ .

*Proof.* See Section 6.5.3. □

Now consider a measurement effect,  $E_n$ , acting on  $\mathcal{H}_{\mathcal{M}}^{\otimes n}$ , i.e. a Hermitian operator  $E_n \in \mathcal{L}(\mathcal{H}_{\mathcal{M}}^{\otimes n})$  such that  $0 \leq E_n \leq \mathbb{1}_{\mathcal{M}}^{\otimes n}$ . If Eq. (6.3.4) is satisfied by some state  $\rho_{\mathcal{M}} \in \mathcal{D}(\mathcal{H}_{\mathcal{M}})$ , and  $\mathrm{Tr}(E_n \rho_{\mathcal{M}}^{\otimes n}) \neq 0$ , then Lemma 6.3.5 implies

$$\sup_{\sigma_{\mathcal{M}} \in \mathcal{C}_{\mathcal{M}}} \frac{\mathrm{Tr}(E_n \sigma_{\mathcal{M}}^{\otimes n})}{\mathrm{Tr}(E_n \rho_{\mathcal{M}}^{\otimes n})} \geq \binom{nm+d_J-1}{nm}^{-1}. \quad (6.3.6)$$

Therefore, to show that every  $\mathcal{M}$ -product state,  $\rho_{\mathcal{M}}$ , eventually violates Eq. (6.3.4), and thus prove Theorem 6.3.4, it suffices to prove the existence of a sequence of measurement effects,  $n \mapsto E_n$ , such that the above ratio of probabilities, as a function of increasing  $n$ , approaches zero faster than  $\binom{nm+d_J-1}{nm}^{-1}$ , thus violating Eq. (6.3.6).

The particular problem of finding a sequence of measurements such that  $\mathrm{Tr}(E_n \rho_{\mathcal{M}}^{\otimes n})$  stays reasonably large while simultaneously minimizing  $\mathrm{Tr}(E_n \sigma_{\mathcal{M}}^{\otimes n})$  for all  $\sigma_{\mathcal{M}}$  distinct from  $\rho_{\mathcal{M}}$  is related to the problems of asymmetric quantum state discrimination and quantum hypothesis testing [HP91; ON05; PBP22; Hay01; Hay02]. Broadly speaking, existing results in these fields are sufficiently strong to establish the claimed violation of Eq. (6.3.6) for each unrealizable  $\rho_{\mathcal{M}} \notin \mathcal{C}_{\mathcal{M}}$ .

Our specific approach, largely inspired by applications of the spectral estimation technique [KW01] to the spectral QMP [CM06; CHM07; CSW18], relies on ideas developed by Keyl [Key06] for the purposes of quantum state estimation. In the interest of being constructive and non-asymptotic, in Section 6.5.4 it is shown how to construct, for each  $\rho_{\mathcal{M}}$ , an explicit sequence of projection operators,  $n \mapsto E_n$ , such that for all  $n \in \mathbb{N}$ ,

$$\sup_{\sigma_{\mathcal{M}} \in \mathcal{C}_{\mathcal{M}}} \frac{\mathrm{Tr}(E_n \sigma_{\mathcal{M}}^{\otimes n})}{\mathrm{Tr}(E_n \rho_{\mathcal{M}}^{\otimes n})} \leq \exp(-(n - d_{\mathcal{M}}^2) \Omega(\rho_{\mathcal{M}}) + c(\rho_{\mathcal{M}})) \quad (6.3.7)$$

where  $c(\rho_{\mathcal{M}}) \geq 0$  and  $\Omega(\rho_{\mathcal{M}}) \geq 0$  are quantities independent of  $n$  (but dependent on  $\rho_{\mathcal{M}}$ ) and  $d_{\mathcal{M}} = \dim(\mathcal{H}_{\mathcal{M}})$ . Additionally, it is shown that the exponential rate,  $\Omega(\rho_{\mathcal{M}})$ , vanishes if and only if  $\rho_{\mathcal{M}}$  is realizable and thus its positivity can serve as a witness of the unrealizability of  $\rho_{\mathcal{M}}$ .

*Proof of Theorem 6.3.4.* The discussion preceding Theorem 6.3.4 has already established the “only if” portion of Theorem 6.3.4: applying  $\mathrm{Tr}_{mJ \setminus \mathcal{M}}^{\otimes n}$  to Eq. (6.3.2) when  $k = nm$  yields

$$\rho_{\mathcal{M}}^{\otimes n} = \mathrm{Tr}_{mJ \setminus \mathcal{M}}^{\otimes n}(\psi_J^{\otimes nm}) \leq \mathrm{Tr}_{mJ \setminus \mathcal{M}}^{\otimes n}(\Pi_J^{(nm)}). \quad (6.3.8)$$

Therefore, all that remains is to prove the “if” portion of Theorem 6.3.4. Suppose  $\rho_{\mathcal{M}} \in \mathcal{D}(\mathcal{H}_{\mathcal{M}})$  is a state that satisfies Eq. (6.3.4) for some particular value of  $n$ . By combining

Eq. (6.3.7) with Eq. (6.3.6), we conclude

$$\Omega(\rho_{\mathcal{M}}) \leq \frac{\log \binom{nm+d_J-1}{nm} + c(\rho_{\mathcal{M}})}{n - d_{\mathcal{M}}^2}. \quad (6.3.9)$$

Since  $\binom{nm+d_J-1}{nm} \in O(n^{d_J-1})$  is polynomial of degree  $d_J - 1$  in  $n$ , the upper bound above approaches zero in the limit as  $n \rightarrow \infty$ . For finite  $n$ , the inequality in Eq. (6.3.9) merely implies that  $\Omega(\rho_{\mathcal{M}})$  must be small. For the purposes of Theorem 6.3.4, if a state  $\rho_{\mathcal{M}}$  satisfies Eq. (6.3.4) for all  $n$ , Eq. (6.3.9) implies that  $\Omega(\rho_{\mathcal{M}}) = 0$  and thus  $\rho_{\mathcal{M}} \in \mathcal{C}_{\mathcal{M}}$  must be a realizable  $\mathcal{M}$ -product state.  $\square$

## 6.4 Conclusion

This chapter makes progress toward an analytic solution to the quantum marginal problem (QMP) by constructing a countably infinite family of necessary operator inequalities whose satisfaction by a given tuple of density operators is sufficient to conclude their realizability. The primary advantage of this approach is its generality: for any finite Hilbert space dimension(s) and any number of subsystems with arbitrary overlap, the corresponding family of necessary inequalities is shown to be sufficient. The results of this chapter, therefore, constitute the first solution to the QMP for overlapping marginal contexts that is free of existential quantifiers. However, the characterization of realizable density operators produced by this approach is not finite, and thus inherently more challenging to compute. Evidently, further insights will be required to produce a finite set of necessary and sufficient conditions for the overlapping QMP.

## 6.5 Supporting results

### 6.5.1 Schur-Weyl Decompositions

Given any representation of the symmetric group  $\Sigma_n$  over a finite-dimensional complex space, such as the aforementioned tensor-permutation representation  $T : \Sigma_n \rightarrow \mathcal{L}((\mathbb{C}^d)^{\otimes n})$ , Maschke's Theorem guarantees the representation is *completely reducible* and therefore decomposes into irreducible subrepresentations [Sag13, Theorem 1.5.3]. Furthermore, the complete set of non-isomorphic irreducible representations of  $\Sigma_n$  is isomorphic to the set of conjugacy classes of  $\Sigma_n$  [Sag13, Proposition 1.10.1] which itself is isomorphic to the set of partitions of  $n$ .

**Definition 6.5.1.** A *partition of  $n$* ,  $\lambda = (\lambda_1, \dots, \lambda_\ell)$ , is a sequence of non-increasing ( $\lambda_i \geq \lambda_{i+1}$ ) positive integers ( $\lambda_i \in \mathbb{N}$ ) whose total sum is  $n$  ( $\sum_i \lambda_i = n$ ). The length of  $\lambda$  is denoted by  $\ell = \ell(\lambda)$ . The *set of all partitions of  $n$*  will be denoted  $\mathbb{Y}_n$ .

For each partition  $\lambda \in \mathbb{Y}_n$ , let  $\omega_\lambda : \Sigma_n \rightarrow \mathcal{V}_\lambda$  denote the corresponding irreducible representation of  $\Sigma_n$ , otherwise known as the Specht module for  $\lambda$  [Sag13, Section 2.3].



Using this notation, the Maschke decomposition of the tensor-permutation representation  $T : \Sigma_n \rightarrow \mathcal{L}((\mathbb{C}^d)^{\otimes n})$  yields a decomposition of  $(\mathbb{C}^d)^{\otimes n}$ ,

$$(\mathbb{C}^d)^{\otimes n} \cong \bigoplus_{\lambda \in \mathbb{Y}_n} \mathcal{V}_\lambda \otimes \mathcal{M}_\lambda^d, \quad (6.5.1)$$

where the  $\mathcal{M}_\lambda^d$  denotes the *multiplicity space*, whose dimension counts the number of isomorphic copies of  $\mathcal{V}_\lambda$  in  $(\mathbb{C}^d)^{\otimes n}$ . It is also worth noting that  $\dim(\mathcal{M}_\lambda^d) > 0$  if and only if  $\ell(\lambda) \leq d$  [Sag13] and therefore the above summands over  $\lambda$  are implicitly restricted to the subset  $\mathbb{Y}_n^d \subseteq \mathbb{Y}_n$  of partitions of  $n$  with length at most  $d$ . Another result, referred to as Schur-Weyl duality [Pro07, Chapter 9], implies that the multiplicity space  $\mathcal{M}_\lambda^d$  itself supports an irreducible representation of  $\mathrm{GL}(d)$ , denoted  $\pi_\lambda : \mathrm{GL}(d) \rightarrow \mathcal{L}(\mathcal{M}_\lambda^d)$ . Let  $|\phi_\lambda\rangle \in \mathcal{M}_\lambda^d$  denote the unique highest weight vector of  $\mathcal{M}_\lambda^d$  characterized by the property that

$$\pi_\lambda(\mathrm{diag}(x_1, \dots, x_d)) |\phi_\lambda\rangle = \prod_{i=1}^d x_i^{\lambda_i} |\phi_\lambda\rangle \quad (6.5.2)$$

for all  $\mathrm{diag}(x_1, \dots, x_d) \in \mathrm{GL}(d)$ . Furthermore, for each partition  $\lambda \in \mathbb{Y}_n^d$ , let

$$\iota_\lambda : \mathcal{V}_\lambda \otimes \mathcal{M}_\lambda^d \hookrightarrow (\mathbb{C}^d)^{\otimes n} \quad (6.5.3)$$

be the  $\Sigma_n \times \mathrm{GL}(d)$ -intertwining isometry from the isotypic subspace  $\mathcal{V}_\lambda \otimes \mathcal{M}_\lambda^d$  into  $(\mathbb{C}^d)^{\otimes n}$  associated to  $\lambda$ . Furthermore, let

$$\Pi_d^\lambda := (\iota_\lambda)(\iota_\lambda)^\dagger \quad (6.5.4)$$

denote the corresponding orthogonal projection operator acting on  $(\mathbb{C}^d)^{\otimes n}$ .

**Proposition 6.5.2.** *Let  $Q \in \mathcal{L}((\mathbb{C}^d)^{\otimes n})$  be an  $\Sigma_n$ -invariant operator in the sense that*

$$\forall g \in \Sigma_n : T(g)QT^\dagger(g) = Q. \quad (6.5.5)$$

*Then  $Q$  admits of the following decomposition,*

$$Q = \bigoplus_{\lambda \in \mathbb{Y}_n^d} \mathbb{1}_{\mathcal{V}_\lambda} \otimes \tau_\lambda(Q), \quad (6.5.6)$$

*where the  $\lambda$ -component of  $Q$ ,  $\tau_\lambda(Q) \in \mathcal{L}(\mathcal{M}_\lambda^d)$ , is defined as*

$$\tau_\lambda(Q) = \frac{\mathrm{Tr}_{\mathcal{V}_\lambda}(\iota_\lambda^\dagger Q \iota_\lambda)}{\dim(\mathcal{V}_\lambda)}. \quad (6.5.7)$$

*Proof.* This follows from an application of Schur's lemma [Hal15, Theorem 4.29].  $\square$

**Definition 6.5.3.** Let  $\pi_\lambda : \mathrm{GL}(d) \rightarrow \mathcal{M}_\lambda^d$  be the irreducible representation of  $\mathrm{GL}(d)$  with highest weight vector  $|\phi_\lambda\rangle \in \mathcal{M}_\lambda^d$  where  $\lambda \in \mathbb{Y}_n^d$ . For each unitary operator  $U \in \mathrm{U}(d) \subseteq \mathrm{GL}(d)$ , let the *twirled highest weight vector* be defined as  $|\phi_\lambda^U\rangle := \pi_\lambda(U) |\phi_\lambda\rangle \in \mathcal{M}_\lambda^d$ .

Consider the quantity  $\langle \phi_\lambda^U | \tau_\lambda(\rho^{\otimes n}) | \phi_\lambda^U \rangle$ , which depends only on  $\lambda \in \mathbb{Y}_n^d$  and  $U^\dagger \rho U \in \mathcal{D}(\mathbb{C}^d)$ . In Section 6.5.4, specifically Proposition 6.5.12, we will show that this quantity can be extended to a formula that remains well-defined even when  $\lambda$  is permitted to be a non-increasing sequence of *non-negative real numbers*.

## 6.5.2 Spectra & Partitions

The purpose of this subsection is to develop a connection between i) partitions  $\lambda \in \mathbb{Y}_n^d$  with length at most  $d$ , and ii) the possible eigenvalues of density operators  $\rho \in \mathcal{D}(\mathbb{C}^d)$ .

**Definition 6.5.4.** A subset  $C \subseteq \mathbb{R}^d$  is called a *convex cone* if it is closed under

- i) addition: for any  $x, y \in C$ ,  $x + y \in C$ , and
- ii) multiplication: for any  $x \in C$ , and  $a \geq 0$ ,  $ax \in C$ .

Two convex cones that are relevant here will be the cone of non-negative real numbers,

$$\mathbb{R}_{\geq 0}^d = \{(x_1, \dots, x_d) \in \mathbb{R}^d \mid \forall i : x_i \geq 0\}, \quad (6.5.8)$$

and the subset of non-increasing non-negative real numbers,

$$\mathbb{R}_{\geq 0}^{d;\downarrow} = \{(x_1, \dots, x_d) \in \mathbb{R}^d \mid x_1 \geq \dots \geq x_d \geq 0\}. \quad (6.5.9)$$

While there is a natural surjective map from  $\mathbb{R}_{\geq 0}^d$  to  $\mathbb{R}_{\geq 0}^{d;\downarrow}$  which sorts the elements of  $(x_1, \dots, x_d)$  in a non-increasing order, there is also a *bijective* linear map  $\gamma : \mathbb{R}_{\geq 0}^d \rightarrow \mathbb{R}_{\geq 0}^{d;\downarrow}$  which takes partial sums. Specifically,  $\gamma$  maps  $y = (y_1, \dots, y_d) \in \mathbb{R}_{\geq 0}^d$  to  $\gamma(y) = (\gamma_1(y), \dots, \gamma_d(y))$  where

$$\gamma_i(y) = y_i + y_{i+1} + \dots + y_d. \quad (6.5.10)$$

The inverse of  $\gamma$ , henceforth denoted  $\delta : \mathbb{R}_{\geq 0}^{d;\downarrow} \rightarrow \mathbb{R}_{\geq 0}^d$ , takes finite differences; specifically,  $\delta$  maps  $x = (x_1, \dots, x_d) \in \mathbb{R}_{\geq 0}^{d;\downarrow}$  to  $\delta(x) = (\delta_1(x), \dots, \delta_d(x))$ , where

$$\delta_i(x) = \delta_i(x_1, \dots, x_k) = \begin{cases} x_i - x_{i+1} & 1 \leq i < d \\ x_d & i = d \end{cases}. \quad (6.5.11)$$

**Definition 6.5.5.** For each  $x = (x_1, \dots, x_d) \in \mathbb{R}_{\geq 0}^d$ , the *size of  $x$* ,  $|x|$ , is the sum of its elements

$$|x| = x_1 + \dots + x_d, \quad (6.5.12)$$

If  $x$  is not equal to all-zero  $d$ -tuple,  $x \neq (0, \dots, 0)$ , then  $|x| > 0$ , and the *normalization of  $x$*  is defined as

$$\frac{x}{|x|} = \left( \frac{x_1}{|x|}, \dots, \frac{x_d}{|x|} \right). \quad (6.5.13)$$

Two subsets of  $\mathbb{R}_{\geq 0}^{d;\downarrow}$  will be crucial to the results of [Section 6.5.4](#). The first subset was already discussed in [Section 6.5.1](#), namely partitions of  $n$  with length at most  $d$ :  $\mathbb{Y}_n^d \subseteq \mathbb{R}_{\geq 0}^{d;\downarrow}$ . If the length,  $\ell$ , of  $\lambda = (\lambda_1, \dots, \lambda_\ell)$  is strictly less than  $d$ , then it can be viewed as a element of  $\mathbb{R}_{\geq 0}^{d;\downarrow}$  by padding  $\lambda$  with  $d - \ell$  zeros, i.e.  $\lambda \cong (\lambda_1, \dots, \lambda_\ell, 0, \dots, 0)$ . The second subset corresponds to the set of possible eigenvalues, or spectra, of density operators  $\mathcal{D}(\mathbb{C}^d)$ .

**Definition 6.5.6.** The set of *spectra*, or sorted probability distributions, is

$$\mathbb{S}^d = \{s \in \mathbb{R}_{\geq 0}^{d;\downarrow} \mid \sum_{i=1}^d s_i = |s| = 1\}. \quad (6.5.14)$$



While the normalization of any partition,  $\lambda \in \mathbb{Y}_n^d$ , is a spectrum,  $\frac{\lambda}{n} \in \mathbb{S}^d$ , multiplying a spectrum,  $s \in \mathbb{S}^d$ , by  $n \in \mathbb{N}$  does not necessarily produce a partition because the entries of  $ns$ ,  $(ns_1, \dots, ns_d)$ , may not be integer-valued. Nevertheless,  $ns \in \mathbb{R}_{\geq 0}^{d \downarrow}$  can always be approximated by a partition,  $\lambda \in \mathbb{Y}_n^d$ , so that  $|\lambda_i - ns_i| \leq 1$  for all  $i \in \{1, \dots, d\}$ <sup>1</sup>. In [Section 6.5.4](#), it will be useful to consider approximating  $ns$  with a partition,  $\lambda$ , in a different manner, where i) degeneracies of  $s$  are preserved, i.e.,  $\delta_i(s) = 0 \implies \delta_i(\lambda) = 0$ , and ii) non-degeneracies of  $s$  are adequately represented, e.g.,  $\delta_i(\lambda) \geq \delta_i(ns)$ . The next lemma shows that this can always be accomplished by partitions,  $\lambda$ , whose size is approximately  $n$ .

**Proposition 6.5.7.** *Let  $s = (s_1, \dots, s_k) \in \mathbb{S}_d$  be a spectrum and  $n \in \mathbb{N}$ . Let*

$$\lambda = (\lambda_1, \dots, \lambda_d) \in \mathbb{N}_{\geq 0}^{d \downarrow} \quad (6.5.15)$$

be defined by

$$\lambda_i = \lceil n(s_i - s_{i+1}) \rceil + \dots + \lceil n(s_{d-1} - s_d) \rceil + \lceil ns_d \rceil, \quad (6.5.16)$$

such that  $\delta_i(\lambda) = \lceil \delta_i(ns) \rceil$  holds. Then  $\lambda$  is a partition of size  $|\lambda|$  where

$$n \leq |\lambda| \leq n + \binom{d+1}{2} - 1. \quad (6.5.17)$$

*Proof.* First note that for all  $1 \leq i \leq d$ ,

$$\epsilon_i := \lceil \delta_i(ns) \rceil - \delta_i(ns), \quad (6.5.18)$$

is upper and lower bounded by  $0 \leq \epsilon_i < 1$ . From this observation, it will be shown that  $\lambda$  approximates  $ns$ , specifically,

$$0 \leq \lambda_i - ns_i < d - i + 1, \quad (6.5.19)$$

To prove [Eq. \(6.5.19\)](#), we use (reverse) induction starting from the base case of  $i = d$ . Since  $\lambda_d = \delta_d(\lambda) = \lceil \delta_d(ns) \rceil = \lceil ns_d \rceil = ns_d + \epsilon_d$ , we have  $\lambda_d - ns_d = \epsilon_d$  and thus [Eq. \(6.5.19\)](#) holds when  $i = d$ . Then, assuming [Eq. \(6.5.19\)](#) holds for  $i = j + 1$ , we prove it holds for  $i = j$ . Since

$$\delta_j(\lambda) = \lambda_j - \lambda_{j+1} = \lceil \delta_j(ns) \rceil \quad (6.5.20)$$

$$= \delta_j(ns) + \epsilon_j = ns_j - ns_{j+1} + \epsilon_j, \quad (6.5.21)$$

we conclude that  $\lambda_j - ns_j = \lambda_{j+1} - ns_{j+1} + \epsilon_j$  and thus  $0 \leq \lambda_j - ns_j < d - j + 1$  which is [Eq. \(6.5.19\)](#) for  $i = j$ . Finally, [Eq. \(6.5.17\)](#) follows from [Eq. \(6.5.19\)](#) by summing over all  $i$ :

$$0 \leq |\lambda| - n|s| < \sum_{i=1}^d (d - i + 1) = \binom{d+1}{2}. \quad (6.5.22)$$

Since  $|\lambda|$  is necessarily an integer,  $|s| = 1$ , and the upper bound above is strict, [Eq. \(6.5.19\)](#) holds.  $\square$

The bounds proven above are also tight for every  $d$ : if  $s = \binom{d+1}{2}^{-1}(d, d-1, \dots, 1)$ , then  $n = 1$  or  $n = \binom{d+1}{2}$  yields  $\lambda = (d, d-1, \dots, 1)$  with size  $|\lambda| = \binom{d+1}{2}$  which achieves the upper bound when  $n = 1$  and the lower bound when  $n = \binom{d+1}{2}$ .

<sup>1</sup>An explicit scheme for accomplishing such an approximation is to let  $t = n - \sum_i \lceil ns_i \rceil$  and define  $\lambda_i = \lceil ns_i \rceil + 1$  whenever  $i \leq t$  and  $\lambda_i = \lceil ns_i \rceil$  whenever  $i > t$ .

### 6.5.3 Proof of Lemma 6.3.5

*Proof of Lemma 6.3.5.* Let  $d_J = \dim(\mathcal{H}_J)$ , let  $\mu_J$  be the  $U(d_J)$ -invariant Haar probability measure over the space of pure states  $\mathbb{P}(\mathcal{H}_J)$ , For any  $k \in \mathbb{N}$ , the orthogonal projection operator,  $\Pi_J^{(k)}$ , onto the symmetric subspace,  $\vee^k \mathcal{H}_J$ , is proportional to the expected value of  $\psi_J^{\otimes k}$  when  $\psi_J$  is sampled according to the probability measure  $\mu_J$ :

$$\Pi_J^{(k)} = \binom{k+d_J-1}{k} \int_{\mathbb{P}(\mathcal{H}_J)} \mu_J(d\psi_J) \psi_J^{\otimes k}, \quad (6.5.23)$$

where the normalization factor is simply  $\text{Tr}[\Pi_J^{(k)}] = \binom{k+d_J-1}{k}$ . The proof of Eq. (6.5.23) follows from Schur's lemma (see [Har13, Proposition 6]).

Next, define the map  $\tau_{\mathcal{M}} : \mathbb{P}(\mathcal{H}_J) \rightarrow \mathcal{D}(\mathcal{H}_{\mathcal{M}})$  by

$$\tau_{\mathcal{M}}(\psi_J) = \text{Tr}_{mJ \setminus \mathcal{M}}(\psi_J^{\otimes m}). \quad (6.5.24)$$

Let  $\nu_{\mathcal{M}}$  be the push-forward measure of  $\mu_J$  through  $\tau_{\mathcal{M}}$ , i.e.  $\nu_{\mathcal{M}} = \mu_J \circ \tau_{\mathcal{M}}^{-1}$ .

Next note that the coefficients  $\tau_{\mathcal{M}}(\psi_J)$  are homogeneous polynomials of degree  $m$  in the coefficients of  $\psi_J$ , and thus  $\tau_{\mathcal{M}}$  is continuous and measurable. Additionally, by construction, the image of  $\tau_{\mathcal{M}}$  is precisely the set of realizable  $\mathcal{M}$ -product states  $\mathcal{C}_{\mathcal{M}}$ . Therefore, since  $\mathbb{P}(\mathcal{H}_J)$  is compact (as  $\mathcal{H}_J$  is finite-dimensional),  $\mathcal{C}_{\mathcal{M}}$  is also compact (and thus closed).

Moreover, the support of the pushforward measure,  $\nu_{\mathcal{M}} = \mu_J \circ \tau_{\mathcal{M}}^{-1}$  is equal to  $\mathcal{C}_{\mathcal{M}}$ . This is because, by the closure of  $\mathcal{C}_{\mathcal{M}}$ ,  $\rho_{\mathcal{M}} \notin \mathcal{C}_{\mathcal{M}}$  implies there exists an open set,  $O$  containing  $\rho_{\mathcal{M}}$ , such that  $O \cap \mathcal{C}_{\mathcal{M}} = \emptyset$  which implies  $\nu_{\mathcal{M}}(O) = \mu_J(\tau_{\mathcal{M}}^{-1}(O)) = \mu_J(\emptyset) = 0$ , i.e.  $\rho_{\mathcal{M}}$  is not in support of  $\nu_{\mathcal{M}}$ . Moreover, if  $\sigma_{\mathcal{M}} \in \mathcal{C}_{\mathcal{M}}$  and  $O'$  is any open set containing  $\sigma_{\mathcal{M}}$ ,  $\tau_{\mathcal{M}}^{-1}(O')$  is non-empty and open (by continuity of  $\tau_{\mathcal{M}}$ ) in  $\mathbb{P}(\mathcal{H}_J)$  and thus  $\nu_{\mathcal{M}}(O') = \mu_J(\tau_{\mathcal{M}}^{-1}(O')) > 0$ . Therefore, because  $\nu_{\mathcal{M}}(O') > 0$  for all open sets containing  $\sigma_{\mathcal{M}}$ ,  $\sigma_{\mathcal{M}}$  is in the support  $\nu_{\mathcal{M}}$ .

Finally, using Eq. (6.5.23), linearity of  $\text{Tr}_{mJ \setminus \mathcal{M}}$ , and a change of variables,

$$\begin{aligned} & \text{Tr}_{mJ \setminus \mathcal{M}}^{\otimes n}(\Pi_J^{(nm)}) \\ & \propto \int_{\mathbb{P}(\mathcal{H}_J)} \mu_J(d\psi_J) (\text{Tr}_{mJ \setminus \mathcal{M}}(\psi_J^{\otimes m}))^{\otimes n}, \end{aligned} \quad (6.5.25)$$

$$= \int_{\mathbb{P}(\mathcal{H}_J)} \mu_J(d\psi_J) (\tau(\psi_J))^{\otimes n}, \quad (6.5.26)$$

$$= \int_{\mathcal{C}_{\mathcal{M}}} \nu_{\mathcal{M}}(d\sigma_{\mathcal{M}}) \sigma_{\mathcal{M}}^{\otimes n}. \quad (6.5.27)$$

□

### 6.5.4 Keyl Divergence & State Discrimination

The purpose of this subsection is to prove Theorem 6.5.14 which can be interpreted as an explicit strategy for asymmetric quantum state discrimination. While Theorem 6.5.14 is exclusively used by this paper in the proof of Theorem 6.3.4, it may be of independent interest. Many of the results of this subsection come directly from Keyl's work on a large-deviation-theoretic approach to quantum state estimation [Key06]. The only additional

insight not taken from [Key06] is the use of Proposition 6.5.7 in the proof of Theorem 6.5.14. Both Section 6.5.1 and Section 6.5.2 are considered prerequisites for this subsection.

**Definition 6.5.8.** Let  $x \in \mathbb{R}_{\geq 0}^{d;\downarrow}$  and let  $\rho \in \mathcal{D}(\mathbb{C}^d)$ , define

$$\Delta_x(\rho) = \prod_{i=1}^d \text{lpm}_i(\rho)^{\delta_i(x)}, \quad (6.5.28)$$

where  $\text{lpm}_i(\rho)$  is the  $i$ th (leading) principal minor of  $\rho$ , i.e. the determinant of the upper-left  $i \times i$ -submatrix of  $\rho$  with respect to some fixed, computational basis  $\{|0\rangle, \dots, |d\rangle\}$ .<sup>2</sup>

The function defined in Eq. (6.5.28) is also referred to as the *generalized power function* [OW16, Notation 4.1]. Note however, in [OW16],  $x$  is restricted to be a partition with length at most  $d$ , while  $\rho$  is permitted to be any  $d \times d$  complex-valued matrix.

**Proposition 6.5.9.** Let  $s = (s_1, \dots, s_d) \in \mathbb{S}^d$  be the spectrum of  $\sigma \in \mathcal{D}(\mathbb{C}^d)$ . For all  $x = (x_1, \dots, x_d) \in \mathbb{R}_{\geq 0}^{k;\downarrow}$ ,

$$\Delta_x(\sigma) \leq \Delta_x(\text{diag}(s_1, \dots, s_d)) = \prod_{i=1}^d s_i^{x_i}, \quad (6.5.29)$$

with equality holding if and only if  $\sigma = \text{diag}(s_1, \dots, s_d)$ .

*Proof.* Consider any  $i \in \{1, \dots, d\}$ . Let  $(s_1^{(i)}, \dots, s_i^{(i)})$  with  $s_1^{(i)} \geq \dots \geq s_i^{(i)}$  denote the eigenvalues of the  $i \times i$  leading principal submatrix of  $\sigma$  so that  $\text{lpm}_i(\sigma) = s_1^{(i)} \cdots s_i^{(i)}$ . According to Cauchy's interlacing theorem (see [Fis05] or [HJ85, Thm. 4.3.17] noting the reversed ordering of labels), for all  $1 < i \leq d$ ,

$$s_1^{(i)} \geq s_1^{(i-1)} \geq s_2^{(i)} \geq \dots \geq s_{i-1}^{(i)} \geq s_{i-1}^{(i-1)} \geq s_i^{(i)}. \quad (6.5.30)$$

Therefore, for any  $k$  and  $i$  such that  $1 \leq k \leq i \leq d$ ,  $s_k = s_k^{(d)} \geq s_k^{(i)} \geq s_k^{(k)}$ . Therefore, for all  $i \in \{1, \dots, d\}$ ,

$$\text{lpm}_i(\sigma) \leq s_1 \cdots s_i, \quad (6.5.31)$$

with equality holding (for all  $i$ ) only if  $\sigma = \text{diag}(s_1, \dots, s_d)$ .  $\square$

**Proposition 6.5.10.** Let  $s = (s_1, \dots, s_d) \in \mathbb{S}^d \subseteq \mathbb{R}_{\geq 0}^{d;\downarrow}$ . Then

$$\Delta_s(\text{diag}(s)) = \prod_{i=1}^d s_i^{s_i} = \exp(-H(s)) > 0. \quad (6.5.32)$$

where  $H(s) = -\sum_{i=1}^d s_i \log s_i$  is the Shannon entropy of  $s$ .

**Corollary 6.5.11.** Let  $\rho \in \mathcal{D}(\mathbb{C}^d)$  have spectrum  $s = (s_1, \dots, s_d) \in \mathbb{S}^d$  and let  $U \in \text{U}(d)$  be a unitary such that  $\rho = U \text{diag}(s_1, \dots, s_d) U^\dagger$  and let  $\sigma \in \mathcal{D}(\mathbb{C}^d)$ . Then

$$\frac{\Delta_s(U^\dagger \sigma U)}{\Delta_s(\text{diag}(s))} = \exp(-K(\rho \parallel \sigma)) \quad (6.5.33)$$

where  $K(\rho \parallel \sigma)$  is defined as

$$K(\rho \parallel \sigma) = \sum_{i=1}^d s_i \log s_i - \delta_i(s) \log \text{lpm}_i(U^\dagger \sigma U), \quad (6.5.34)$$

where  $K(\rho \parallel \sigma) \in [0, \infty]$  and  $K(\rho \parallel \sigma) = 0$  if and only if  $\sigma = \rho$ .

---

<sup>2</sup>If it happens that  $\text{lpm}_i(\rho) = 0$  and  $\delta_i(x) = 0$  for some index  $i$ , then the indeterminate expression  $0^0$  is taken to be equal to 1.

Notice that if  $\rho$  and  $\sigma$  are simultaneously diagonalized by  $U$  so that

$$U^\dagger \sigma U = \text{diag}(t_1, \dots, t_d), \quad (6.5.35)$$

then the quantity  $K(\rho \parallel \sigma)$  simplifies to the classical relative entropy,  $D(s \parallel t) = \sum_{i=1}^d s_i (\log s_i - \log t_i)$ , also known as Kullback-Liebler divergence [Kul97]. Also note that, in general,  $K(\rho \parallel \sigma)$  does not equal the quantum relative entropy  $S(\rho \parallel \sigma) = \text{Tr}(\rho(\log \rho - \log \sigma))$ , but is nevertheless bounded by  $K(\rho \parallel \sigma) \leq S(\rho \parallel \sigma)$  [Key06]. For these reasons, we refer to the quantity  $K(\rho \parallel \sigma)$  as *Keyl-divergence*.

**Proposition 6.5.12.** *Let  $\lambda \in \mathbb{Y}_n^d$  be a partition of  $n \in \mathbb{N}$  and let  $|\phi_\lambda^U\rangle \in \mathcal{M}_\lambda^d$  be the twirled highest weight vector for  $U \in \text{U}(d)$  (see Definition 6.5.3). Then for all  $\sigma \in \mathcal{D}(\mathbb{C}^d)$ ,*

$$\langle \phi_\lambda^U | \tau_\lambda(\sigma^{\otimes |\lambda|}) | \phi_\lambda^U \rangle = \Delta_\lambda(U^\dagger \sigma U). \quad (6.5.36)$$

*Proof.* This result is noted by Keyl as [Key06, Eqs. (141) & (151)] with reference to [Zhe73, Sec. 49].  $\square$

Henceforth, define the projection operator  $\Phi_\lambda^U \in \mathcal{L}(H^{\otimes |\lambda|})$  by

$$\Phi_\lambda^U := \iota_\lambda(|\phi_\lambda^U\rangle \langle \phi_\lambda^U| \otimes \mathbb{1}_{\mathcal{V}_\lambda}) \iota_\lambda^\dagger, \quad (6.5.37)$$

so that,

$$\text{Tr}(\Phi_\lambda^U \rho^{\otimes |\lambda|}) = \dim(\mathcal{V}_\lambda) \Delta_\lambda(U^\dagger \rho U). \quad (6.5.38)$$

**Corollary 6.5.13.** *Let  $\rho, \sigma \in \mathcal{D}(\mathbb{C}^d)$  and let  $U \in \text{U}(d)$  diagonalize  $\rho$ ,*

$$\rho = U \text{diag}(s_1, \dots, s_d) U^\dagger. \quad (6.5.39)$$

*If  $\rho$  has rational spectra  $s = (s_1, \dots, s_d) \in \mathbb{S}^d$ , i.e. there exists a  $q \in \mathbb{N}$  such that  $qs \in \mathbb{Y}_q^d$  is a partition of  $q$ , then for all  $n \in \mathbb{N}$ ,*

$$\frac{\text{Tr}(\Phi_{nqs}^U \sigma^{\otimes nq})}{\text{Tr}(\Phi_{nqs}^U \rho^{\otimes nq})} = \exp(-nqK(\rho \parallel \sigma)). \quad (6.5.40)$$

*Proof.* The proof follows from Eq. (6.5.38) and Corollary 6.5.11. When Eq. (6.5.38) is applied to the numerator and denominator on the left-hand-side of Eq. (6.5.40), the common factor of  $\dim(\mathcal{V}_{nqs}) > 0$  cancels out.  $\square$

A result similar to Corollary 6.5.13 holds for arbitrary states  $\rho \in \mathcal{D}(\mathbb{C}^d)$ , e.g., for states that do not have rational spectra.

**Theorem 6.5.14.** *Let  $\rho, \sigma \in \mathcal{D}(\mathbb{C}^d)$ , let  $s = (s_1, \dots, s_d) \in \mathbb{S}^d$  be the spectrum of  $\rho$ , and let  $U \in \text{U}(d)$  diagonalize  $\rho$ , such that  $\rho = U \text{diag}(s_1, \dots, s_d) U^\dagger$ . Then there exists a sequence,  $n \mapsto \lambda^n \in \mathbb{Y}_n^d$  of partitions such that for all  $n \in \mathbb{N}$ ,*

$$\frac{\text{Tr}(\Phi_{\lambda^n}^U \sigma^{\otimes n})}{\text{Tr}(\Phi_{\lambda^n}^U \rho^{\otimes n})} \leq D(s) \exp(-(n - \binom{d+1}{2} + 1)K(\rho \parallel \sigma)). \quad (6.5.41)$$

where  $D(s)$  is a constant depending only on  $s$ .

*Proof.* The proof relies on an explicit construction of a sequence,  $n \mapsto \lambda^n$ , that satisfies the claim. For each non-negative integer  $k \in \mathbb{N}_{\geq 0}$ , let  $\mu^k$  be the partition characterized by  $\delta_i(\mu^k) = \lceil \delta_i(k s) \rceil$  (see [Proposition 6.5.7](#)). Then for any state  $\eta \in \mathcal{D}(\mathbb{C}^d)$ , we claim

$$\Delta_{ks}(\eta)\Delta_{\mu^1}(\eta) \leq \Delta_{\mu^k}(\eta) \leq \Delta_{ks}(\eta). \quad (6.5.42)$$

To see the upper bound, note that for all  $i \in \{1, \dots, d\}$ ,  $\delta_i(\mu^k) \geq \delta_i(k s) > 0$  so  $\text{lpm}_i(\eta)^{\delta_i(\mu^k)} \leq \text{lpm}_i(\eta)^{\delta_i(k s)}$  since  $\text{lpm}_i(\eta) < 1$  (noting [Eq. \(6.5.31\)](#)). For the lower bound, note that  $\delta_i(\mu^k) = \lceil \delta_i(k s) \rceil \leq \delta_i(k s) + \lceil \delta_i(s) \rceil = \delta_i(k s) + \delta_i(\mu^1)$ , so  $\Delta_{\mu^k}(\eta) \geq \Delta_{ks}(\eta)\Delta_{\mu^1}(\eta)$  holds. Next, apply [Corollary 6.5.11](#) and [Eq. \(6.5.42\)](#) (the upper bound when  $\eta = U^\dagger \sigma U$ , and the lower bound when  $\eta = U^\dagger \rho U = \text{diag}(s_1, \dots, s_d)$ ) to obtain

$$\frac{\text{Tr}(\Phi_{\mu^k}^U \sigma^{\otimes |\mu^k|})}{\text{Tr}(\Phi_{\mu^k}^U \rho^{\otimes |\mu^k|})} \leq \frac{\exp(-kK(\rho \parallel \sigma))}{\Delta_{\mu^1}(\text{diag}(s_1, \dots, s_d))}. \quad (6.5.43)$$

Now, notice that [Eq. \(6.5.43\)](#) is almost in the form of [Eq. \(6.5.41\)](#). The main obstacle remaining is simply that the size of  $\mu^k$  needs to be decoupled from the spectra of  $\rho$ . Fortunately, [Proposition 6.5.7](#) guarantees that  $|\mu^k|$  is, at least, approximately equal to  $k$  because  $k \leq |\mu^k| \leq k + \binom{d+1}{2} - 1$ . Moreover, since  $|\mu^{k+1}| \geq |\mu^k|$ , there always exists at least one value of  $k$  such that  $\mu^k$  has size approximately  $n$  for any  $n \in \mathbb{N}$ ; specifically, there exists a  $k \in \mathbb{N}$  such that

$$n - \binom{d+1}{2} + 1 \leq |\mu^k| \leq n. \quad (6.5.44)$$

Now simply define  $\lambda^n \in \mathbb{Y}_n^d$  by

$$\lambda^n = (\mu_1^k + n - |\mu^k|, \mu_2^k, \dots, \mu_d^k), \quad (6.5.45)$$

where  $k$  is the largest such that  $\mu_k$  satisfies [Eq. \(6.5.44\)](#). Note that when  $n$  is small ( $n < \binom{d+1}{2} - 1$ ), is entirely possible for  $k = 0$  and  $\mu_0 = (0, 0, \dots, 0)$ , in which case,  $\lambda^n = (n, 0, \dots, 0)$ . This definition ensures

$$\text{Tr}(\Phi_{\lambda^n}^U \rho^{\otimes n}) = s_1^{n-|\mu^k|} \text{Tr}(\Phi_{\mu^k}^U \rho^{\otimes |\mu^k|}), \quad (6.5.46)$$

$$\text{Tr}(\Phi_{\lambda^n}^U \sigma^{\otimes n}) \leq \text{Tr}(\Phi_{\mu^k}^U \sigma^{\otimes |\mu^k|}). \quad (6.5.47)$$

Therefore, from [Eqs. \(6.5.43\)](#) and [\(6.5.44\)](#), we conclude [Eq. \(6.5.41\)](#) where  $D(s)$  is the constant

$$D(s) = s_1^{1-\binom{d+1}{2}} (\Delta_{\mu^1}(\text{diag}(s_1, \dots, s_d)))^{-1} \quad (6.5.48)$$

$$= s_1^{1-\binom{d+1}{2}} \prod_{i=1}^d (s_1 s_2 \dots s_i)^{-\lceil \delta_i(s) \rceil}. \quad (6.5.49)$$

□

**Remark 6.5.15.** To derive the inequality in [Eq. \(6.3.7\)](#) from the result of [Theorem 6.5.14](#), note that  $\binom{d+1}{2} - 1 \leq d^2$  and substitute

- i)  $E_n = \Phi_{\lambda^n}^U$ ,
- ii)  $\Omega(\rho_{\mathcal{M}}) = \inf_{\sigma_{\mathcal{M}} \in \mathcal{C}_{\mathcal{M}}} K(\rho_{\mathcal{M}} \parallel \sigma_{\mathcal{M}})$ ,
- iii)  $c(\rho_{\mathcal{M}}) = \log D(\text{spec}(\rho_{\mathcal{M}}))$ , and
- iv)  $d = d_{\mathcal{M}} = \dim(\mathcal{H}_{\mathcal{M}})$ .

The claim that  $\Omega(\rho_{\mathcal{M}})$  vanishes if and only if  $\rho_{\mathcal{M}} \in \mathcal{C}_{\mathcal{M}}$  follows from the compactness of  $\mathcal{C}_{\mathcal{M}}$  and the following corollary.

**Corollary 6.5.16.** *Let  $\mathcal{C} \subseteq \mathcal{D}(\mathbb{C}^d)$  be compact. Define*

$$\Omega(\rho) := \inf_{\sigma \in \mathcal{C}} K(\rho \parallel \sigma). \quad (6.5.50)$$

*Then  $\Omega(\rho) = 0$  if and only if  $\rho \in \mathcal{C}$ .*

*Proof.* If  $\rho \in \mathcal{C}$ , then [Corollary 6.5.11](#) implies  $\Omega(\rho) = K(\rho \parallel \rho) = 0$ . Otherwise if  $\rho \notin \mathcal{C}$ , then consider, for each fixed  $x \in \mathbb{R}_{\geq 0}^{d \downarrow}$  and  $U \in \text{U}(d)$ , that the function  $\sigma \mapsto \Delta_x(U^\dagger \sigma U) \in [0, 1]$  is continuous as  $\text{lpm}_i(U^\dagger \sigma U)$  is a polynomial in the coefficients of  $\sigma$ . The compactness of  $\mathcal{C}$  guarantees (using the extreme value theorem) the supremum is attained by some  $\sigma_x^U \in \mathcal{C}$ :

$$\Delta_x(U^\dagger \sigma_x^U U) = \sup_{\sigma \in \mathcal{C}} \Delta_x(U^\dagger \sigma U). \quad (6.5.51)$$

Since  $\Omega(\rho) = K(\rho \parallel \sigma_x^U)$  and  $\rho \neq \sigma_x^U$ , we conclude, from [Corollary 6.5.11](#), that  $\Omega(\rho) \neq 0$ .  $\square$

## 6.6 Special Cases

### 6.6.1 Pure vs. Full QMP

One might wonder why the version of the QMP considered in this paper ([Problem 6.2.1](#)) seems to be exclusively interested in the existence of joint states that are *pure*,  $\psi_J \in \mathbb{P}(\mathcal{H}_J)$ , instead of the more general *density operator*,  $\rho_J \in \mathcal{D}(\mathcal{H}_J)$ . In order to distinguish between these two types of QMP, the former is sometimes called the *pure* QMP, while the latter is sometimes called the *mixed* QMP. Of these two variants, the mixed QMP is arguably a much closer analogy to the classical marginals problem [[FC12](#)].

The distinction between these two variants is strongest when the marginal scenario under consideration,  $\mathcal{M} = (S_1, \dots, S_m)$ , has disjoint marginal contexts, i.e.  $S_i \cap S_j = \emptyset$  for all  $i \neq j$ , or equivalently  $\mathcal{H}_{\mathcal{M}} = \mathcal{H}_{S_1} \otimes \dots \otimes \mathcal{H}_{S_m} \cong \mathcal{H}_J$ . Under the assumption of disjoint marginal contexts, the mixed QMP becomes trivial; every collection of density operators  $(\rho_{S_1}, \dots, \rho_{S_m})$  are the  $\mathcal{M}$ -marginals of the density operator  $\rho_J = \rho_{S_1} \otimes \dots \otimes \rho_{S_m}$ . On the other hand, under this assumption, the pure QMP remains non-trivial.

At the level of generality considered in this paper, wherein the marginal scenario  $\mathcal{M} = (S_1, \dots, S_m)$  is permitted to contain overlapping marginal contexts, e.g.,  $S_i \cap S_j \neq \emptyset$ , the distinction becomes less important because the marginals of a mixed state can equivalently be viewed as the marginals of any of its purifications. Specifically, there exists a density operator  $\sigma_J \in \mathcal{D}(\mathcal{H}_J)$  with marginals  $\rho_S = \text{Tr}_{J \setminus S}(\sigma_J)$  for all  $S \in \mathcal{M}$  if and only if there exists

a joint pure state  $\psi_{JJ'} \in \mathbb{P}(\mathcal{H}_J \otimes \mathcal{H}_{J'})$  (where  $\mathcal{H}_{J'} \cong \mathcal{H}_J$ ) such that  $(\text{Tr}_{J \setminus S} \otimes \text{Tr}_{J'}) (\sigma_J) = \rho_S$  for all  $S \in \mathcal{M}$ . Consequently, the techniques developed in this paper, which directly apply to the pure QMP, can also be applied to any instance of the mixed QMP without substantial modification.

## 6.6.2 Diagrammatics

The purpose of this subsection is to briefly introduce some diagrammatic notation that will be useful for performing a few calculations in [Section 6.6.3](#). The particular notations involving symmetrization and antisymmetrization used here ([Eq. \(6.6.11\)](#) and onward), are taken from Cvitanović's excellent textbook [[Cvi08](#)] on diagrammatic calculations of invariants of Lie groups, and are essentially the same those used by Penrose [[Pen71](#)]. For a categorical justification of this notation, see [[Sel12](#)]. For further applications within quantum theory, see [[CK10](#); [WBC11](#); [BB17](#)].

The essential idea is to depict linear operators,  $L : \mathcal{H}_X \rightarrow \mathcal{H}_Y$ , by pictures with corresponding inputs and outputs:

$$\begin{array}{c} \downarrow Y \\ \boxed{L} \\ \uparrow X \end{array} . \quad (6.6.1)$$

Of course, two linear operators can be combined in at least three different ways; specifically, by addition  $+$ , tensor product  $\otimes$ , and sequential composition  $\circ$ . These operations are depicted respectively as

$$\begin{array}{c} \downarrow Y \\ \boxed{L+M} \\ \uparrow X \end{array} = \begin{array}{c} \downarrow Y \\ \boxed{L} \\ \uparrow X \end{array} + \begin{array}{c} \downarrow Y \\ \boxed{M} \\ \uparrow X \end{array} , \quad (6.6.2)$$

$$\begin{array}{c} \downarrow Y \downarrow Z \\ \boxed{L \otimes M} \\ \uparrow X \uparrow W \end{array} = \begin{array}{c} \downarrow Y \\ \boxed{L} \\ \uparrow X \end{array} \begin{array}{c} \downarrow Z \\ \boxed{M} \\ \uparrow W \end{array} , \quad (6.6.3)$$

$$\begin{array}{c} \downarrow Z \\ \boxed{M \circ L} \\ \uparrow X \end{array} = \begin{array}{c} \downarrow Z \\ \boxed{M} \\ \downarrow Y \\ \boxed{L} \\ \uparrow X \end{array} . \quad (6.6.4)$$

Important special cases of this notation include the identity operator  $\mathbb{1}_X : \mathcal{H}_X \rightarrow \mathcal{H}_X$ ,

$$\begin{array}{c} \downarrow X \\ \boxed{I} \\ \uparrow X \end{array} = \left| \uparrow X \right. , \quad (6.6.5)$$

and vectors  $|\psi_X\rangle : \mathbb{C} \rightarrow \mathcal{H}_X$  (and their conjugates  $\langle\psi_X| : \mathcal{H}_X \rightarrow \mathbb{C}$ ) as

$$\begin{array}{c} \downarrow X \\ \boxed{\psi} \end{array} , \quad \text{and} \quad \begin{array}{c} \boxed{\psi} \\ \uparrow X \end{array} . \quad (6.6.6)$$

This notation is especially elegant for depicting two concepts frequently encountered in this paper: the partial trace and permutations.

First, given a bipartite operator  $L : \mathcal{H}_X \otimes \mathcal{H}_Y \rightarrow \mathcal{H}_X \otimes \mathcal{H}_Y$ , the partial trace  $\text{Tr}_Y$  over  $Y$ , is depicted as

$$\begin{array}{c} \downarrow X \\ \boxed{\text{Tr}_Y(L)} \\ \uparrow X \end{array} = \begin{array}{c} X \downarrow \\ \boxed{L} \\ X \uparrow \end{array} \Bigg\} Y . \quad (6.6.7)$$

The trace over the identity operator  $\mathbb{1}_X : \mathcal{H}_X \rightarrow \mathcal{H}_X$ , which is equal to the dimension of  $\mathcal{H}_X$ , is therefore depicted as a closed loop

$$d_X = \dim(\mathcal{H}_X) = \boxed{1}_X = \left. \vphantom{\int} \right\}_X = \bigcirc_X. \quad (6.6.8)$$

Second, the tensor permutation representation,  $T_X : \Sigma_m \rightarrow \mathcal{L}(\mathcal{H}_X^{\otimes m})$ , of the symmetric group,  $\Sigma_m$ , has elements depicted naturally as follows. When  $m = 2$ ,  $\Sigma_2 = \{e, (12)\}$ , and the identity  $T_X(e)$  and swap  $T_X((12))$  are depicted respectively by

$$\begin{array}{c} \downarrow \\ x \end{array} \begin{array}{c} \downarrow \\ x \end{array}, \quad \text{and} \quad \begin{array}{c} \diagdown \quad \diagup \\ x \quad x \end{array}. \quad (6.6.9)$$

Analogously, for  $m = 3$ , the  $3! = 6$  permutations in  $\Sigma_3$  are depicted by

$$\begin{array}{ccc} \begin{array}{c} \downarrow \\ x \end{array} \begin{array}{c} \downarrow \\ x \end{array} \begin{array}{c} \downarrow \\ x \end{array}, & \begin{array}{c} \diagdown \quad \diagup \\ x \quad x \end{array} \begin{array}{c} \downarrow \\ x \end{array}, & \begin{array}{c} \downarrow \\ x \end{array} \begin{array}{c} \diagdown \quad \diagup \\ x \quad x \end{array}, \\ \begin{array}{c} \diagdown \quad \diagup \\ x \quad x \end{array} \begin{array}{c} \downarrow \\ x \end{array}, & \begin{array}{c} \diagdown \quad \diagup \\ x \quad x \end{array} \begin{array}{c} \diagdown \quad \diagup \\ x \quad x \end{array}, & \begin{array}{c} \diagdown \quad \diagup \\ x \quad x \end{array} \begin{array}{c} \diagdown \quad \diagup \\ x \quad x \end{array}. \end{array} \quad (6.6.10)$$

The orthogonal projection operator,  $\Pi_X^{(2)}$ , onto the symmetric subspace  $\vee^2 \mathcal{H}_X \subseteq \mathcal{H}_X^{\otimes 2}$ , referred to as the *symmetrization operator*, is given the following unique notation:

$$\begin{array}{c} \downarrow \downarrow \\ \hline \downarrow \downarrow \\ x \quad x \end{array} := \begin{array}{c} \downarrow \downarrow \\ \boxed{\Pi_X^{(2)}} \\ \downarrow \downarrow \\ x \quad x \end{array} = \frac{1}{2} \begin{array}{c} \downarrow \\ x \end{array} \begin{array}{c} \downarrow \\ x \end{array} + \frac{1}{2} \begin{array}{c} \diagdown \quad \diagup \\ x \quad x \end{array}. \quad (6.6.11)$$

Similarly, the orthogonal projection operator onto the antisymmetric subspace  $\wedge^2 \mathcal{H}_X \subseteq \mathcal{H}_X^{\otimes 2}$ , denoted by  $\Pi_X^{(1,1)}$  and referred to as the *antisymmetrization operator*, is depicted in a complementary manner:

$$\begin{array}{c} \downarrow \downarrow \\ \hline \downarrow \downarrow \\ x \quad x \end{array} := \begin{array}{c} \downarrow \downarrow \\ \boxed{\Pi_X^{(1,1)}} \\ \downarrow \downarrow \\ x \quad x \end{array} = \frac{1}{2} \begin{array}{c} \downarrow \\ x \end{array} \begin{array}{c} \downarrow \\ x \end{array} - \frac{1}{2} \begin{array}{c} \diagdown \quad \diagup \\ x \quad x \end{array}. \quad (6.6.12)$$

Generalizing this notation for the orthogonal projection operators onto the symmetric and antisymmetric subspaces of  $\mathcal{H}_X^{\otimes m}$  for  $m > 2$  can be done recursively as follows. For the sake of clarity, the Hilbert space label,  $X$ , can often be omitted without introducing ambiguity.

$$\begin{array}{c} \downarrow \downarrow \dots \downarrow \\ \hline \downarrow \downarrow \dots \downarrow \end{array} = \frac{1}{m} \left( \begin{array}{c} \downarrow \downarrow \dots \downarrow \\ \downarrow \downarrow \dots \downarrow \end{array} + (m-1) \begin{array}{c} \downarrow \downarrow \dots \downarrow \\ \downarrow \downarrow \dots \downarrow \end{array} \right), \quad (6.6.13)$$

$$\begin{array}{c} \downarrow \downarrow \dots \downarrow \\ \hline \downarrow \downarrow \dots \downarrow \end{array} = \frac{1}{m} \left( \begin{array}{c} \downarrow \downarrow \dots \downarrow \\ \downarrow \downarrow \dots \downarrow \end{array} - (m-1) \begin{array}{c} \downarrow \downarrow \dots \downarrow \\ \downarrow \downarrow \dots \downarrow \end{array} \right). \quad (6.6.14)$$

Finally, in order to generalize the above symmetrization and antisymmetrization notation to the case of multipartite Hilbert spaces, e.g.,  $\mathcal{H}_{XY} = \mathcal{H}_X \otimes \mathcal{H}_Y$ , we introduce the following notational definition for the joint symmetrization  $\Pi_{XY}^{(2)}$ :

$$\begin{array}{c} \downarrow \downarrow \\ \hline \downarrow \downarrow \\ x \quad x \quad y \quad y \end{array} = \frac{1}{2} \begin{array}{c} \downarrow \downarrow \\ x \quad x \end{array} \begin{array}{c} \downarrow \downarrow \\ y \quad y \end{array} + \frac{1}{2} \begin{array}{c} \diagdown \quad \diagup \\ x \quad x \end{array} \begin{array}{c} \diagdown \quad \diagup \\ y \quad y \end{array}. \quad (6.6.15)$$



### 6.6.3 The degree one case

This subsection explores the strength of the constraint imposed by Eq. (6.3.4) for the special case when  $n = 1$  (equivalently Eq. (6.3.3)) for the purposes of detecting unrealizable  $\mathcal{M}$ -product states. For marginal scenarios involving disjoint marginal contexts, it will be shown that Eq. (6.3.3) happens to be satisfied by *all*  $\mathcal{M}$ -product states, and therefore is useless for the QMP. For at least some marginal scenarios involving non-disjoint marginal contexts, it will be shown that Eq. (6.3.3) is already capable of witnessing the unrealizability of certain  $\mathcal{M}$ -product states. Finally, it is shown that for some (admittedly degenerate) marginal scenarios, the constraint imposed by Eq. (6.3.3) is also sufficient for the corresponding QMP.

Throughout this subsection, unipartite subsystems are labeled alphabetically, e.g.,  $A, B, C \dots$ , and their respective dimensions denoted by lower-case letters, e.g.,  $a = d_A, b = d_B$ , etc.

Consider the marginal scenario  $\mathcal{M} = (A, B)$  ( $m = 2$ ) for the joint context  $J = AB$ . In this scenario, the QMP is already fully solved:  $\rho_A$  and  $\rho_B$  are the marginals of some pure state  $\psi_{AB}$  if and only if  $\text{spec}(\rho_A) = \text{spec}(\rho_B)$  (see Section 6.6.4). To what extent, if any, does Eq. (6.3.3) reproduce this known solution? If  $\rho_A$  and  $\rho_B$  are the marginals of some pure state  $\psi_{AB}$ , Eq. (6.3.3) implies

$$\rho_A \otimes \rho_B \leq (\text{Tr}_B \otimes \text{Tr}_A)(\Pi_{AB}^{(2)}). \quad (6.6.16)$$

To calculate the right-hand side of the above inequality, it will be convenient to use the diagrammatic notation introduced in Section 6.6.2. Specifically,  $\Pi_{AB}^{(2)}$  can be depicted using Eq. (6.6.15) (with  $X, Y$  substituted by  $A, B$ ), and thus  $(\text{Tr}_B \otimes \text{Tr}_A)(\Pi_{AB}^{(2)})$  is equal to

$$\begin{array}{c} \uparrow \downarrow \\ \text{---} \text{---} \\ \uparrow \downarrow \end{array} \begin{array}{c} \uparrow \downarrow \\ \text{---} \text{---} \\ \uparrow \downarrow \end{array} = \frac{1}{2} \begin{array}{c} \uparrow \downarrow \\ \text{---} \text{---} \\ \uparrow \downarrow \end{array} \begin{array}{c} \uparrow \downarrow \\ \text{---} \text{---} \\ \uparrow \downarrow \end{array} + \frac{1}{2} \begin{array}{c} \uparrow \downarrow \\ \text{---} \text{---} \\ \uparrow \downarrow \end{array} \begin{array}{c} \uparrow \downarrow \\ \text{---} \text{---} \\ \uparrow \downarrow \end{array}, \quad (6.6.17)$$

$$= \frac{ab}{2} \begin{array}{c} \uparrow \downarrow \\ \text{---} \text{---} \\ \uparrow \downarrow \end{array} \begin{array}{c} \uparrow \downarrow \\ \text{---} \text{---} \\ \uparrow \downarrow \end{array} + \frac{1}{2} \begin{array}{c} \uparrow \downarrow \\ \text{---} \text{---} \\ \uparrow \downarrow \end{array} \begin{array}{c} \uparrow \downarrow \\ \text{---} \text{---} \\ \uparrow \downarrow \end{array}, \quad (6.6.18)$$

$$= \frac{1 + ab}{2} \begin{array}{c} \uparrow \downarrow \\ \text{---} \text{---} \\ \uparrow \downarrow \end{array}. \quad (6.6.19)$$

Therefore, Eq. (6.6.16) is equivalent to  $\rho_A \otimes \rho_B \leq \frac{1}{2}(1 + ab)\mathbb{1}_A \otimes \mathbb{1}_B$ , i.e.

$$\begin{array}{c} \uparrow \\ \boxed{\rho_A} \\ \uparrow_A \end{array} \begin{array}{c} \uparrow \\ \boxed{\rho_B} \\ \uparrow_B \end{array} \leq \frac{1 + ab}{2} \begin{array}{c} \uparrow \\ \text{---} \\ \uparrow_A \end{array} \begin{array}{c} \uparrow \\ \text{---} \\ \uparrow_B \end{array}, \quad (6.6.20)$$

which is an inequality satisfied by all  $(A, B)$ -product states  $\rho_A \otimes \rho_B$  because  $\rho_X \leq \mathbb{1}_X$  already holds for all  $\rho_X \in \mathcal{D}(\mathcal{H}_X)$  and  $ab \geq 1$ . In fact, it is not too difficult to show that when  $\mathcal{M} = (X_1, \dots, X_k)$  contains disjoint contexts, i.e.  $X_i \cap X_j = \emptyset$  for  $i \neq j$ , the inequality in Eq. (6.3.3) is always trivial because  $\text{Tr}_{m\mathcal{J}\setminus\mathcal{M}}(T_J(\pi)) \geq \mathbb{1}_{\mathcal{M}}$  holds for all  $\pi \in \Sigma_k$  and thus  $\text{Tr}_{m\mathcal{J}\setminus\mathcal{M}}(\Pi_J^{(m)}) \geq \mathbb{1}_{\mathcal{M}}$  also. Fortunately, the same is not necessarily true for non-disjoint marginal scenarios.

For an example of a non-trivial instance of Eq. (6.3.3), consider the marginal scenario  $\mathcal{M} = (AB, AC, BC)$  for the joint context  $J = ABC$ . In this scenario, the aforementioned

operator inequality becomes

$$\rho_{AB} \otimes \rho_{AC} \otimes \rho_{BC} \leq (\text{Tr}_C \otimes \text{Tr}_B \otimes \text{Tr}_A)(\Pi_{ABC}^{(3)}). \quad (6.6.21)$$

The projector  $\Pi_{ABC}^{(3)}$  onto  $\vee^3(\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C)$  can be expressed as

$$\begin{aligned} \begin{array}{ccc} \uparrow\uparrow\uparrow & \uparrow\uparrow\uparrow & \uparrow\uparrow\uparrow \\ \downarrow\downarrow\downarrow & \downarrow\downarrow\downarrow & \downarrow\downarrow\downarrow \\ A & B & C \end{array} - \begin{array}{ccc} \uparrow\uparrow\uparrow & \uparrow\uparrow\uparrow & \uparrow\uparrow\uparrow \\ \downarrow\downarrow\downarrow & \downarrow\downarrow\downarrow & \downarrow\downarrow\downarrow \\ A & C & B \end{array} - \begin{array}{ccc} \uparrow\uparrow\uparrow & \uparrow\uparrow\uparrow & \uparrow\uparrow\uparrow \\ \downarrow\downarrow\downarrow & \downarrow\downarrow\downarrow & \downarrow\downarrow\downarrow \\ A & B & C \end{array} = \frac{1}{3!} \left[ \begin{array}{ccc} ||| & ||| & ||| \\ A & B & C \end{array} + \begin{array}{ccc} \times & \times & \times \\ A & B & C \end{array} \right] + \\ \left[ \begin{array}{ccc} \times & \times & \times \\ A & B & C \end{array} + \begin{array}{ccc} \times & \times & \times \\ A & C & B \end{array} + \begin{array}{ccc} \times & \times & \times \\ A & B & C \end{array} + \begin{array}{ccc} \times & \times & \times \\ A & C & B \end{array} + \begin{array}{ccc} \times & \times & \times \\ A & B & C \end{array} + \begin{array}{ccc} \times & \times & \times \\ A & C & B \end{array} \right]. \end{aligned} \quad (6.6.22)$$

Therefore,  $(\text{Tr}_C \otimes \text{Tr}_B \otimes \text{Tr}_A)(\Pi_{ABC}^{(3)})$  becomes

$$\begin{aligned} \begin{array}{ccc} \uparrow\uparrow\uparrow & \uparrow\uparrow\uparrow & \uparrow\uparrow\uparrow \\ \downarrow\downarrow\downarrow & \downarrow\downarrow\downarrow & \downarrow\downarrow\downarrow \\ A & B & C \end{array} = \frac{1}{3!} \left[ abc \begin{array}{ccc} ||| & ||| & ||| \\ A & B & C \end{array} + \right. \\ \left. + c \begin{array}{ccc} ||| & \times & ||| \\ A & B & C \end{array} + a \begin{array}{ccc} \times & ||| & ||| \\ A & B & C \end{array} + b \begin{array}{ccc} ||| & \times & ||| \\ A & B & C \end{array} + 2 \begin{array}{ccc} \times & \times & \times \\ A & B & C \end{array} \right]. \end{aligned} \quad (6.6.23)$$

To show that Eq. (6.6.21) is a non-trivial constraint, we consider consider the case of three qubits, i.e.  $a = b = c = 2$ . For a given pair of qubits, the unique antisymmetric pure state (also called the singlet state),  $\Phi = |\Phi\rangle\langle\Phi| \in \mathbb{P}(\mathbb{C}^2 \otimes \mathbb{C}^2)$ , can be identified with  $|\Phi\rangle = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle)$  and depicted as follows

$$\begin{array}{c} \uparrow\downarrow \\ \downarrow\uparrow \\ \Phi \end{array} = \begin{array}{c} \uparrow\downarrow \\ \downarrow\uparrow \end{array}, \quad \text{s.t.} \quad \begin{array}{c} \times \\ \times \end{array} = - \begin{array}{c} \uparrow\downarrow \\ \downarrow\uparrow \end{array}, \quad \begin{array}{c} \uparrow\uparrow \\ \downarrow\downarrow \end{array} = 1. \quad (6.6.24)$$

When applied to  $(\text{Tr}_C \otimes \text{Tr}_B \otimes \text{Tr}_A)(\Pi_{ABC}^{(3)})$  (assuming  $a = b = c = 2$ ) we obtain the identity

$$\begin{array}{ccc} \begin{array}{c} \uparrow\uparrow \\ \downarrow\downarrow \end{array} & \begin{array}{c} \uparrow\uparrow \\ \downarrow\downarrow \end{array} & \begin{array}{c} \uparrow\uparrow \\ \downarrow\downarrow \end{array} \\ \begin{array}{c} \uparrow\downarrow \\ \downarrow\uparrow \end{array} & \begin{array}{c} \uparrow\downarrow \\ \downarrow\uparrow \end{array} & \begin{array}{c} \uparrow\downarrow \\ \downarrow\uparrow \end{array} \end{array} = \frac{1}{3!}(abc - a - b - c - 2) = 0, \quad (6.6.25)$$

which, when combined with Eq. (6.6.21) proves that the  $(AB, AC, BC)$  marginals of a three-qubit pure state  $\psi_{ABC}$  always satisfy

$$\begin{array}{c} \begin{array}{ccc} \uparrow\uparrow & \uparrow\uparrow & \uparrow\uparrow \\ \downarrow\downarrow & \downarrow\downarrow & \downarrow\downarrow \\ \rho_{AB} & \rho_{BC} & \rho_{CA} \end{array} \\ \begin{array}{ccc} \uparrow\downarrow & \uparrow\downarrow & \uparrow\downarrow \\ \downarrow\uparrow & \downarrow\uparrow & \downarrow\uparrow \end{array} \end{array} = 0. \quad (6.6.26)$$

An example of an unrealizable triple of states  $(\rho_{AB}, \rho_{AC}, \rho_{BC})$  whose unrealizability is witnessed by the above equality constraint is the triple of anti-correlated states,  $\rho_{AB} = \rho_{AC} = \rho_{BC} = \frac{1}{2}(|01\rangle\langle 01| + |10\rangle\langle 10|)$ , where the left-hand side evaluates to  $2^{-5}$ . Other examples includes the triple of singlets  $\rho_{AB} = \rho_{AC} = \rho_{BC} = \Phi$  (with value  $2^{-4}$ ), or the triple of maximally mixed states  $\rho_{AB} = \rho_{AC} = \rho_{BC} = \frac{I}{2} \otimes \frac{I}{2}$  (with value  $2^{-6}$ ). An example of an inconsistent triple of states for which Eq. (6.6.26) happens to be satisfied is  $\rho_{AB} = \rho_{BC} = |00\rangle\langle 00|$  and  $\rho_{AC} = |11\rangle\langle 11|$ .

To conclude, consider the rather non-standard marginal scenario  $\mathcal{M} = (X, X)$  for the joint context  $J = X$ . Taken literally, the QMP for this marginal scenario is to determine, for any given pair of states  $\rho_X, \sigma_X \in \mathcal{D}(\mathcal{H}_X)$ , whether or not there exists a pure state

$\psi_X \in \mathbb{P}(\mathcal{H}_X)$  such that  $\rho_X = \psi_X$  and  $\sigma_X = \psi_X$ . This marginal scenario can be regarded as “non-standard” for at least two reasons: (i) the marginal context  $X$  is repeated twice in  $\mathcal{M}$ , and (ii) since  $X$  is not a proper subset of  $X$ ,  $\rho_X$  and  $\sigma_X$  are not proper marginals of  $\psi_X$ . Taken together, the QMP for this scenario has a simple solution:  $\rho_X$  and  $\sigma_X$  are realizable if and only if they are both pure states and equal to each other. Nevertheless, in this scenario Eq. (6.3.3) is a valid constraint; in particular, it simplifies to  $\rho_X \otimes \sigma_X \leq \Pi_X^{(2)}$ , or diagrammatically

$$\begin{array}{c} \downarrow \\ \boxed{\rho_X} \\ \downarrow \\ X \end{array} \begin{array}{c} \downarrow \\ \boxed{\sigma_X} \\ \downarrow \\ X \end{array} \leq \begin{array}{c} \downarrow \downarrow \\ \boxed{\phantom{\rho_X \sigma_X}} \\ \downarrow \downarrow \\ X \quad X \end{array}. \quad (6.6.27)$$

The above inequality implies that  $\text{Tr}(\sigma_X \rho_X) = 1$  since

$$0 \leq \begin{array}{c} \text{---} \\ \boxed{\rho_X} \quad \boxed{\sigma_X} \end{array} \leq \begin{array}{c} \text{---} \\ \boxed{\phantom{\rho_X \sigma_X}} \end{array} = 0, \quad (6.6.28)$$

and

$$\begin{array}{c} \text{---} \\ \boxed{\rho_X} \quad \boxed{\sigma_X} \end{array} = \frac{1}{2} \begin{array}{c} \downarrow \\ \boxed{\rho_X} \\ \downarrow \\ X \end{array} \begin{array}{c} \downarrow \\ \boxed{\sigma_X} \\ \downarrow \\ X \end{array} - \frac{1}{2} \begin{array}{c} \downarrow \downarrow \\ \boxed{\phantom{\rho_X \sigma_X}} \\ \downarrow \downarrow \\ X \quad X \end{array}, \quad (6.6.29)$$

$$= \frac{1}{2} \left( 1 - \begin{array}{c} \downarrow \\ \boxed{\sigma_X} \\ \downarrow \\ \boxed{\rho_X} \\ \downarrow \\ X \end{array} \right). \quad (6.6.30)$$

Since  $\text{Tr}(\sigma_X \rho_X) = 1$  holds if and only if  $\sigma_X = \rho_X = \psi_X$  for some pure state  $\psi_X$ , we see that Eq. (6.3.3), which becomes Eq. (6.6.27), is both necessary and *sufficient* for the  $\mathcal{M} = (X, X)$  instance of the QMP.

## 6.6.4 The Bipartite QMP

This subsection considers the bipartite marginal scenario,  $\mathcal{M} = (A, B)$ , for the joint context  $J = AB$ . For this scenario, the QMP is fully solved and admits of a simple solution:  $\rho_A$  and  $\rho_B$  are the marginals of a joint pure state  $\psi_{AB} \in \mathbb{P}(\mathcal{H}_{AB})$  if and only if they have the same spectrum [TV15; Kly04]. A natural question arises: how does Theorem 6.3.4 recover this well-known result?

To answer this question, first let  $a = \dim(\mathcal{H}_A)$ ,  $b = \dim(\mathcal{H}_B)$ , and let  $\ell = \min(a, b)$ . For this scenario, Eq. (6.3.4) becomes

$$(\rho_A \otimes \rho_B)^{\otimes n} \leq (\text{Tr}_B \otimes \text{Tr}_A)^{\otimes n} (\Pi_{AB}^{(2n)}). \quad (6.6.31)$$

Now let  $s_A \in \mathbb{S}^a$  and  $s_B \in \mathbb{S}^b$  be the spectra of  $\rho_A$  and  $\rho_B$ . Using the results of Section 6.5.4 (or essentially the spectral estimation theorem [KW01; CM06]), together with Eq. (6.5.23), it is possible to show that the exponential factor in Eq. (6.3.7),  $\Omega(\rho_A \otimes \rho_B)$ , depends only on  $r_A$  and  $r_B$  and is equal to:

$$\Omega(\rho_A \otimes \rho_B) = \inf_{r \in \mathbb{S}^\ell} (D(s_A \| r) + D(s_B \| r)), \quad (6.6.32)$$

where  $D(p \| q)$  is the relative entropy  $D(p \| q) = \sum_i p_i (\log p_i - \log q_i)$ . Since  $D(p \| q)$  only vanishes if  $p = q$ ,  $\Omega(\rho_A \otimes \rho_B)$  only vanishes if  $s_A = s_B$ . Therefore, we conclude that  $\rho_A$

and  $\rho_B$  are the  $(A, B)$ -marginals of a pure state  $\psi_{AB} \in \mathbb{P}(\mathbb{C}^a \otimes \mathbb{C}^b)$  if and only if they have equal spectra. Additionally, using Pinsker's inequality [RW09],  $\|p - q\|_1^2 \leq 2D(p\|q)$ , and the triangle inequality for  $\|\cdot\|_1$ , we obtain:

$$\|s_A - s_B\|_1^2 \leq 3(\|s_A - r\|_1^2 + \|s_B - r\|_1^2) \quad (6.6.33)$$

$$\leq 6(D(s_A\|r) + D(s_B\|r)). \quad (6.6.34)$$

Therefore,  $\Omega(\rho_A \otimes \rho_B) \geq \|s_A - s_B\|_1^2/6$ .

A more direct consequence of Eq. (6.6.31) is the following proposition.

**Proposition 6.6.1.** *Let  $n \in \mathbb{N}$  and let  $\alpha \in \mathbb{Y}_n^a$  and  $\beta \in \mathbb{Y}_n^b$  be partitions. If  $\rho_A \otimes \rho_B$  satisfies Eq. (6.6.31), then*

$$s_\alpha(r_A)s_\beta(r_B) \leq \sum_{\lambda \in \mathbb{Y}_{2n}^\ell} c_{\alpha\beta}^\lambda \frac{\dim(\mathcal{M}_\lambda^a) \dim(\mathcal{M}_\lambda^b)}{\dim(\mathcal{V}_\lambda)}, \quad (6.6.35)$$

where  $r_A$  and  $r_B$  are the spectra of  $\rho_A$  and  $\rho_B$  respectively. Additionally,  $s_\alpha$  and  $s_\beta$  are Schur functions [Sra16; Sag13] and  $c_{\alpha\beta}^\lambda$  is the Littlewood-Richardson coefficient [LR34; Ful00; PPY19].

*Proof.* One of the most powerful tools for decomposing bipartite Hilbert spaces, specifically the symmetric subspace of a bipartite system  $\vee^k(\mathcal{H}_A \otimes \mathcal{H}_B) \cong V_{(k)}^{ab}$ , is known as  $\text{GL}(a) \times \text{GL}(b)$ -duality [How87] (see also [Wal14, Eq. (2.25)]):

$$V_{(k)}^{ab} \cong \bigoplus_{\lambda \in \mathbb{Y}_k^\ell} \mathcal{M}_\lambda^a \otimes \mathcal{M}_\lambda^b, \quad (6.6.36)$$

where  $\ell = \min(a, b)$ . Using this result, and applying  $\Pi_A^\alpha \otimes \Pi_B^\beta$  to the right-hand-side of Eq. (6.3.4), we obtain

$$\begin{aligned} & \text{Tr}_{AB}^{\otimes n} \{ (\Pi_A^\alpha \otimes \Pi_B^\beta) (\text{Tr}_B^{\otimes n} \otimes \text{Tr}_A^{\otimes n}) (\Pi_{AB}^{(2n)}) \} \\ &= \dim(\mathcal{V}_\alpha) \dim(\mathcal{V}_\beta) \sum_{\lambda \in \mathbb{Y}_{2n}^\ell} c_{\alpha\beta}^\lambda \frac{\dim(\mathcal{M}_\lambda^a) \dim(\mathcal{M}_\lambda^b)}{\dim(\mathcal{V}_\lambda)}, \end{aligned} \quad (6.6.37)$$

where  $c_{\alpha\beta}^\lambda$  counts the multiplicity of the  $\Sigma_n \times \Sigma_n$  irreducible representation space  $\mathcal{V}_\alpha \otimes \mathcal{V}_\beta$  inside  $\mathcal{V}_\lambda$  under the restriction of  $\Sigma_{2n}$  to  $\Sigma_n \times \Sigma_n$ . By comparison, applying  $\Pi_A^\alpha \otimes \Pi_B^\beta$  to the left-hand-side of Eq. (6.3.4) yields

$$\text{Tr}((\Pi_A^\alpha \otimes \Pi_B^\beta)(\rho_A^{\otimes n} \otimes \rho_B^{\otimes n})) \quad (6.6.38)$$

$$= \text{Tr}(\Pi_A^\alpha \rho_A^{\otimes n}) \text{Tr}(\Pi_B^\beta \rho_B^{\otimes n}), \quad (6.6.39)$$

$$= s_\alpha(r_A) \dim(\mathcal{V}_\alpha) s_\beta(r_B) \dim(\mathcal{V}_\beta). \quad (6.6.40)$$

Therefore, Eq. (6.6.31) implies Eq. (6.6.35) and thus the claim holds.  $\square$

### 6.6.5 Fermionic & Bosonic QMP

Our sufficient family of necessary inequality constraints can be modified to handle the fermionic and bosonic variants of the QMP. Recall that a state describing a system of  $p$  fermions (resp. bosons) with  $f$  internal degrees of freedom, is typically modeled by an element of the antisymmetric subspace  $\wedge^p \mathbb{C}^f$  (resp. the symmetric subspace  $\vee^p \mathbb{C}^f$ ). Since  $\wedge^p \mathbb{C}^f$  (resp.  $\vee^p \mathbb{C}^f$ ) can be viewed as a subspace of a  $p$ -partite composite Hilbert space  $(\mathbb{C}^f)^{\otimes p}$ , and  $\vee^n \wedge^p \mathbb{C}^f$  (resp.  $\vee^n \vee^p \mathbb{C}^f$ ) serves as the representation space for an irreducible representation of  $SU(\wedge^p \mathbb{C}^f) \cong SU(\dim(\wedge^p \mathbb{C}^f)) \cong SU(\binom{f}{p})$  (resp.  $SU(\vee^p \mathbb{C}^f) \cong SU(\binom{p+f-1}{p})$ ), the analogue of Eq. (6.5.23) holds and thus an analogue of Lemma 6.3.5 also holds. Altogether, a generalization of Theorem 6.3.4 holds:

**Corollary 6.6.2.** *Let  $\mathcal{H}_V \subseteq \mathcal{H}_J$  be a subspace of a joint Hilbert space,  $\mathcal{H}_J$ . An  $\mathcal{M}$ -product state,  $\rho_{\mathcal{M}} = \rho_{S_1} \otimes \cdots \otimes \rho_{S_m}$ , is realizable by a joint pure state  $\psi_V \in \mathcal{H}_V \subseteq \mathcal{H}_J$  in the subspace  $\mathcal{H}_V$  if and only if for all  $n \in \mathbb{N}$ ,*

$$\rho_{\mathcal{M}}^{\otimes n} \leq \text{Tr}_{mJ \setminus \mathcal{M}}^{\otimes n}(\Pi_V^{(nm)}), \quad (6.6.41)$$

where  $\Pi_V^{(nm)}$  is the projection operator onto the  $nm$ -symmetric subspace  $\vee^{nm} \mathcal{H}_V \subseteq \mathcal{H}_J^{\otimes nm}$ .

### 6.6.6 Counting Solutions to the QMP

Whenever a given  $\mathcal{M}$ -product state,  $\rho_{\mathcal{M}} = \rho_{S_1} \otimes \cdots \otimes \rho_{S_m}$ , is shown to be realizable, a natural follow-up problem is to determine whether or not the joint state,  $\psi_J$ , satisfying Eq. (6.2.9) is unique. For the bipartite marginal scenario,  $\mathcal{M} = (\mathcal{A}, \mathcal{B})$ , if the common spectrum of  $\rho_A$  and  $\rho_B$  is  $s = (s_1, \dots, s_r, 0, \dots, 0)$ , with positive values distinct, i.e.  $s_1 > \cdots > s_r > 0$ , then the *unique* pure state,  $\psi_{AB} \in \mathbb{P}(\mathcal{H}_J)$ , satisfying Eq. (6.2.9) is  $\psi_{AB} = |\psi_{AB}\rangle \langle \psi_{AB}|$  where

$$|\psi_{AB}\rangle = \sum_{i=1}^r \sqrt{s_i} |\phi_A^{(i)}\rangle \otimes |\phi_B^{(i)}\rangle, \quad (6.6.42)$$

where  $\{\phi_A^{(i)}\}_{i=1}^r$  and  $\{\phi_B^{(i)}\}_{i=1}^r$  are the eigenvectors of  $\rho_A$  and  $\rho_B$ . If, however, the common spectrum,  $(s_1, \dots, s_r, 0, \dots, 0)$ , is degenerate in the sense that some of its values are identical, then the solution to Eq. (6.2.9) may not be unique. A familiar example of this phenomenon, for the two-qubit Hilbert space  $\mathcal{H}_J \cong \mathbb{C}^2 \otimes \mathbb{C}^2$ , are the four Bell states all sharing the same pair of maximally-mixed, single-qubit marginals,  $(\frac{1}{2}, \frac{1}{2})$ .

The following result generalizes Eq. (6.3.4) by considering the possibility that an  $\mathcal{M}$ -product may be realizable by multiple, orthogonal, joint states.

**Corollary 6.6.3.** *Let  $\rho_{\mathcal{M}} = \rho_{S_1} \otimes \cdots \otimes \rho_{S_m}$  be an  $\mathcal{M}$ -product state and  $\{\psi_J^{(1)}, \dots, \psi_J^{(v)}\}$  be a set of joint pure states, satisfying *i*) for all  $1 \leq j, k \leq v$ ,*

$$\text{Tr}(\psi_J^{(j)} \psi_J^{(k)}) = |\langle \psi_J^{(j)} | \psi_J^{(k)} \rangle|^2 = \delta_{j,k}, \quad (6.6.43)$$

and *ii*) for all  $1 \leq i \leq m$ , and  $1 \leq k \leq v$ ,

$$\rho_{S_i} = \text{Tr}_{J \setminus S_i}(\psi_J^{(k)}). \quad (6.6.44)$$

Then, the following inequality holds:

$$v^{nm} \rho_{\mathcal{M}}^{\otimes n} \leq \sum_{\lambda \in \mathbb{Y}_{nm}^v} \text{Tr}_{mJ \setminus \mathcal{M}}^{\otimes n}(\Pi_J^\lambda). \quad (6.6.45)$$

*Proof.* Let  $P_V$  be the orthogonal projection operator onto the subspace of  $\mathcal{H}_J$  spanned by  $\{\psi_J^{(k)}\}_{k=1}^v$ , i.e.,

$$P_V = \sum_{k=1}^v \psi_J^{(k)}. \quad (6.6.46)$$

Since each  $\psi_J^{(k)}$  has marginals  $(\rho_{S_1}, \dots, \rho_{S_m})$ , we conclude

$$\text{Tr}_{mJ \setminus \mathcal{M}}^{\otimes n}(P_V^{\otimes nm}) = v^{nm} \rho_{\mathcal{M}}^{\otimes n}. \quad (6.6.47)$$

Furthermore,  $P_V^{\otimes nm}$  commutes with  $T_J(\pi)$  for all  $\pi \in \Sigma_{nm}$  and thus commutes with  $\Pi_J^\lambda$  for every  $\lambda \in \mathbb{Y}_{nm}^{d_J}$ . In fact, we obtain

$$P_V^{\otimes nm} = \sum_{\lambda \in \mathbb{Y}_{nm}^v} \Pi_J^\lambda P_V^{\otimes nm} \Pi_J^\lambda \leq \sum_{\lambda \in \mathbb{Y}_{nm}^v} \Pi_J^\lambda, \quad (6.6.48)$$

because i)  $\Pi_J^\lambda P_V^{\otimes nm} = 0$  for all  $\lambda$  with length  $\ell(\lambda) > v$ , and ii)  $P_V \leq \mathbb{1}_J$ . Applying  $\text{Tr}_{mJ \setminus \mathcal{M}}^{\otimes n}$  yields Eq. (6.6.45).  $\square$

Note that Eq. (6.6.45) is equivalent to Eq. (6.3.4) if  $v = 1$ . Also note that while Eq. (6.6.45) is necessary for the existence of  $v$  orthogonal solutions to the QMP, satisfying Eq. (6.6.45) for all  $n$  is generally insufficient (for  $v > 1$ ) to conclude that  $v$  orthogonal solutions to the QMP exist. For example, if  $v = d_J$ , Eq. (6.6.45) simplifies to

$$\rho_{\mathcal{M}}^{\otimes n} \leq \left( \frac{\mathbb{1}_{\mathcal{M}}}{d_{\mathcal{M}}} \right)^{\otimes n}. \quad (6.6.49)$$

The above constraint is evidently satisfied for all  $n \in \mathbb{N}$ , if and only if  $\rho_{\mathcal{M}}$  is the maximally-mixed  $\mathcal{M}$ -product state, i.e.,  $\rho_{\mathcal{M}} = \mathbb{1}_{\mathcal{M}}/d_{\mathcal{M}}$ . However, such states are generally unrealizable [Kly02], e.g., it can be shown that  $(\frac{\mathbb{1}_A}{a}, \frac{\mathbb{1}_B}{b})$  are not the  $(A, B)$ -marginals of *any* pure state,  $\psi_{AB}$ , if  $a \neq b$ .

# Concluding remarks

This thesis considered the notion of a *quantum realizability problem*, which can be understood as any kind of constraint satisfiability problem involving collections of constraints placed on a quantum state. Throughout the thesis, these constraints are freely interpreted as candidate *properties* the states of a quantum system may or may not possess. If the constraints happen to be satisfiable by some quantum state, then the corresponding property values are said to be *realizable*, otherwise they are *unrealizable*.

This thesis makes progress toward a general method for tackling quantum realizability problems based upon insights from quantum property estimation and quantum tomography. In particular, our main contribution is to demonstrate that the realizability of a given collection of property values is encoded in the asymptotics of a corresponding sequence of probabilities indexed by a positive integer  $n$ . These probabilities, in turn, can be interpreted as the expected likelihood for  $n$ -copies of a random quantum state to behave in a manner which is, in a certain sense, indicative of quantum states exhibiting those property values. In the end, this method produces a hierarchy of necessary conditions, in the form a probabilistic bounds, for the realizability of a given collection of property values which convergences to sufficiency as the number of copies,  $n$ , tends to infinity.

Special emphasis is placed upon quantum realizability problems wherein each property under consideration corresponds to the moment map,  $\Omega$ , of a representation of a complex reductive group,  $G$ , acting on an underlying complex finite-dimensional Hilbert space,  $\mathcal{V}$ . In this setting, the moment map evaluated on the subspace spanned by a given unit vector,  $v$ , denoted by  $\Omega([v])$ , serves two purposes: (i) as a non-commutative gradient of the logarithm of the norm of a point in the  $G$ -orbit of  $v$ , and (ii) as the map assigning to each Hermitian operator in the Lie algebra of  $G$ , the corresponding *expectation value* with respect to the pure quantum state associated to the vector  $v$ . It is this bridge between the statistical moments of Hermitian observables and the algebrogeometric features of group orbits which gives rise to the vast majority of the results presented in this thesis.

Specifically, this thesis builds upon two recent and intimately related results. First and foremost is the *strong-duality* result (see [Section 3.3.4](#), [Theorem 3.3.20](#)) due to Franks and Walter [[FW20](#)] for the problem of non-commutative optimization theory [[Bür+19](#)]. The second result upon which this thesis relies is the generalization of Keyl’s large deviation theory approach to quantum state estimation theory [[Key06](#)] to handle the estimation of arbitrary moment maps developed by Botero, Christandl, and Vrana [[BCV21](#)] (see [Section 4.4.1](#)).

These results are ultimately derivable from the *deformed* strong-duality theorem (see [Section 4.3.4](#), [Theorem 4.3.14](#) or [[FW20](#)]) which, for the purposes of this thesis, provides a correspondence between (i) the value,  $\omega$ , of the moment map of the ray spanned by the



unit vector  $v$ , and (ii) the rate of exponential decay of the norm,  $\|Q_n^\omega v^{\otimes n}\|$  of the  $n$ th tensor power  $v^{\otimes n}$  when projected onto a corresponding sequence of subspaces (by the projection operator  $Q_n^\omega$ ) as  $n$  tends to infinity. The square of the latter of these two quantities can, via the Born rule, be interpreted as the probability for a particular outcome to occur when a collective, covariant measurement is performed on  $n$  copies of the quantum state associated to  $v$ . Altogether, these results enable one to distinguish between those quantum states,  $v$ , whose moment map equals  $\omega$  from those quantum states whose moment map does not equal  $\omega$ ; if  $\Omega([v]) \neq \omega$ , then the probability  $\|Q_n^\omega v^{\otimes n}\|^2$  decays to zero at an exponential rate with increasing  $n$ , otherwise  $\Omega([v]) = \omega$  and the probability  $\|Q_n^\omega v^{\otimes n}\|^2$  does not decay at an exponential rate with increasing  $n$ . These insights are then used in [Chapter 5](#) to derive novel solutions to quantum realizability problems involving a finite, by otherwise arbitrary, collection of moment maps.

Perhaps the most significant application for these new tools was to the quantum marginal problem, where in [Chapter 6](#), it was shown how to construct a sufficient hierarchy of necessary operator inequalities for the realizability of any given collection of quantum marginals. Existing attempts to analytically solve the quantum marginal problem have also produced necessary and sufficient conditions, albeit only for marginal scenarios involving disjoint marginal contexts [[Kly06](#)], marginal scenarios with a small degree of overlap [[CSW18](#)] or additionally only for low-dimensional Hilbert spaces [[Che+14](#)]. While numerical methods for solving quantum marginal problems using semidefinite programming already exist [[Hal07](#); [Yu+21](#)], the results presented in [Chapter 6](#) constitute a significant step toward a fully *analytic* solution to the quantum marginal problem.

Beyond the immediate application of these techniques to instances of quantum realizability problems which remain unsolved, such as the existence of absolutely maximally entangled states for large composite quantum systems [[HC13](#)] or the allocation of von Neumann entropies for more than three subsystems [[Pip03](#); [LW05](#)], there exists numerous promising areas for follow-up research which could build upon the estimation-theoretic themes presented by this thesis, or at the very least, could seek to address or alleviate some of its apparent limitations.

For instance, perhaps the most broadly applicable and powerful result of this thesis, [Corollary 5.1.23](#), suggests the need to calculate the asymptotics of quantities of the form appearing in [Section 5.1.5](#), [Eq. \(5.1.66\)](#) as  $n$  tends to infinity. Despite the myriad of symmetries which may be exploited to simplify calculations (see [Section 5.1.5](#)), a direct calculation for large  $n$  remains both numerically and analytically intractable. To address these difficulties, future research could either seek to characterize which terms in the biriffle expansion ([Theorem 5.1.24](#)) dominate in the limit of large  $n$ , or otherwise seek to derive non-trivial bounds analogous to the bounds from [Lemma 5.1.25](#). Alternatively, using techniques from geometric group theory applied to the symmetric group it appears possible to explore various asymptotic regimes wherein Hilbert space dimensions tend to infinity (analogously to Ref. [[DLN20](#)]).

Another limitation of this approach which future research could seek to address is the assumption that the underlying Hilbert space dimension is both finite and known a priori. From the perspective of performing quantum estimation schemes experimentally, the assumption that the dimensionality is known a priori is difficult to justify. Moreover, the assumption of finite dimensionality was crucial in establishing the inclusivity lower-bounds



used in [Section 5.1.2](#), and would prove challenging to generalize under the assumption of infinite dimensional Hilbert spaces. Nevertheless, it may be possible to reinterpret some of the results of this thesis as providing *witnesses* of large dimensionality which could be experimentally tested.

Future research could also seek to better understand how the convergent hierarchies of necessary inequalities for realizability presented here are related to the convergent hierarchies of semidefinite programs used to solve similar feasibility and optimization programs involving constraints placed on quantum states [[Yu+21](#); [BKM21](#); [NPA08](#)]. At the very least, these two approaches appear superficially related; while the mathematical machinery of this thesis relies on a method for *non-commutative optimization* in the sense of Bürgisser et al. wherein the domain of optimization is a non-abelian group [[Bür+19](#)], the convergent hierarchies of semidefinite programs can also be understood as a method for solving *non-commutative optimization* problems, albeit in a different sense, wherein the domain of optimization is a collection of not-necessarily-commuting variables [[PNA10](#)]. Although the degree to which these two approaches are connected remains unclear, resolving these connections would likely be a fruitful area of research.

Finally, it appears possible to use the results of this thesis to also tackle *approximate* realizability problems. Specifically, by using the bound presented in [Proposition 4.3.12](#) for fixed degree  $n$  along with a [Eq. \(5.1.29\)](#) appearing in the proof of [Theorem 5.1.15](#), it becomes possible to place a lower-bound on the supremum of the deformed capacity, which when combined with the quantitative strengthening of the Kempf-Ness theorem obtained by Bürgisser et al. [[Bür+19](#), Thm. 1.17], it appears possible to place an upper-bound on how *close* a given moment map value is to being realizable. By analyzing this series of bounds in rigorous detail (and generalizing to the case of multiple moment maps as in [Section 5.1.4](#)), it seems plausible that one could place an upper bound on the minimal degree needed to decide a given approximate realizability problem as a function of the approximation parameter,  $\varepsilon$ . Of course, future research is needed to verify the details of this proposal and to determine if the minimal degrees needed to decide an  $\varepsilon$ -realizability problem in this manner grows sufficiently slowly with decreasing  $\varepsilon$  as to be practical.

# Bibliography

- [Aar07] Scott Aaronson. “[The learnability of quantum states](#)”. In: *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* 463.2088 (2007), pp. 3089–3114.
- [Aar18] Scott Aaronson. “Shadow tomography of quantum states”. In: *Proceedings of the 50th annual ACM SIGACT symposium on theory of computing*. 2018, pp. 325–338. arXiv: [1711.01053v2](#).
- [AB11] Samson Abramsky and Adam Brandenburger. “[The sheaf-theoretic structure of non-locality and contextuality](#)”. In: *New Journal of Physics* 13.11 (2011), p. 113036.
- [AB99] A Angelow and MC Batoni. “About Heisenberg Uncertainty Relation (by E. Schrödinger)”. In: *Bulg. J. Phys.* 26 (1999), pp. 193–203. arXiv: [quant-ph/9903100](#).
- [Abb+16] Alastair A Abbott et al. “[Tight state-independent uncertainty relations for qubits](#)”. In: *Mathematics* 4.1 (2016), p. 8.
- [Acu+22] Arturo Acuaviva et al. “The minimal canonical form of a tensor network”. In: (2022). arXiv: [2209.14358](#).
- [AFT20] Albert Aloy, Matteo Fadel, and Jordi Tura. *The quantum marginal problem for symmetric states: applications to variational optimization, nonlocality and self-testing*. 2020. arXiv: [2001.04440v1](#).
- [AKG19] Anirudh Acharya, Theodore Kypraios, and Mădălin Guță. “[A comparative study of estimation methods in quantum tomography](#)”. In: *Journal of Physics A: Mathematical and Theoretical* 52.23 (2019), p. 234001.
- [AL70] Huzihiro Araki and Elliott H Lieb. “[Entropy inequalities](#)”. In: *Communications in Mathematical Physics* 18.2 (1970), pp. 160–170.
- [Amé+21a] Carlos Améndola et al. “[Invariant theory and scaling algorithms for maximum likelihood estimation](#)”. In: *SIAM Journal on Applied Algebra and Geometry* 5.2 (2021), pp. 304–337.
- [Amé+21b] Carlos Améndola et al. “[Toric invariant theory for maximum likelihood estimation in log-linear models](#)”. In: *Algebraic Statistics* 12.2 (2021), pp. 187–211.
- [BB17] Jacob Biamonte and Ville Bergholm. *Tensor networks in a nutshell*. 2017. arXiv: [1708.00006v1](#).

- [BBO20] Xavier Bonet-Monroig, Ryan Babbush, and Thomas E O’Brien. “Nearly optimal measurement scheduling for partial tomography of quantum states”. In: *Physical Review X* 10.3 (2020), p. 031064. arXiv: [1908.05628v3](#).
- [BCV21] Alonso Botero, Matthias Christandl, and Péter Vrana. “Large deviation principle for moment map estimation”. In: *Electronic Journal of Probability* 26 (2021), pp. 1–23.
- [BD72] RE Borland and K Dennis. “The conditions on the one-matrix for three-body fermion wavefunctions with one-rank equal to six”. In: *Journal of Physics B: Atomic and Molecular Physics* 5.1 (1972), p. 7.
- [Ben+08] Jonathan Bennett et al. “The Brascamp–Lieb inequalities: finiteness, structure and extremals”. In: *Geometric and Functional Analysis* 17.5 (2008), pp. 1343–1415.
- [BF14] John C Baez and Tobias Fritz. *A Bayesian characterization of relative entropy*. 2014. arXiv: [1402.3067v2](#).
- [Bha97] Rajendra Bhatia. *Matrix analysis*. New York: Springer, 1997.
- [BKM21] Abhishek Bhardwaj, Igor Klep, and Victor Magron. *Noncommutative Polynomial Optimization*. 2021. arXiv: [2109.00786v1](#).
- [Blu10] Robin Blume-Kohout. “Optimal, reliable estimation of quantum states”. In: *New Journal of Physics* 12.4 (2010), p. 043034.
- [Boo12] Adam D Bookatz. *QMA-complete problems*. 2012. arXiv: [1212.6312v3](#).
- [BR19] Paul Busch and Oliver Reardon-Smith. *On quantum uncertainty relations and uncertainty regions*. 2019. arXiv: [1901.03695v2](#).
- [Bra+17] Fernando GSL Brandão et al. “Quantum SDP solvers: Large speed-ups, optimality, and applications to quantum learning”. In: (2017). arXiv: [1710.02581](#).
- [Bra03] Sergey Bravyi. “Requirements for compatibility between local and multipartite quantum states”. In: (2003). arXiv: [quant-ph/0301014](#).
- [Bri87] Michel Brion. “Sur l’image de l’application moment”. In: *Séminaire d’Algèbre Paul Dubreil et Marie-Paule Malliavin: Proceedings, Paris 1986*. Springer. 1987, pp. 177–192.
- [BRV18] Jim Bryan, Zinovy Reichstein, and Mark Van Raamsdonk. “Existence of locally maximally entangled quantum states via geometric invariant theory”. In: *Annales Henri Poincaré*. Vol. 19. Springer. 2018, pp. 2491–2511. arXiv: [1708.01645v2](#).
- [Bry90] Wlodzimierz Bryc. “Large deviations by the asymptotic value method”. In: *Diffusion processes and related problems in analysis* 1 (1990), pp. 447–472.
- [BS00] Arkady Berenstein and Reyer Sjamaar. “Coadjoint orbits, moment polytopes, and the Hilbert-Mumford criterion”. In: *Journal of the American Mathematical Society* 13.2 (2000), pp. 433–466.
- [BSS06] Paul Butterley, Anthony Sudbery, and Jason Szulc. “Compatibility of subsystem states”. In: *Foundations of Physics* 36.1 (2006), pp. 83–101.

- [BSW23] Mario Berta, David Sutter, and Michael Walter. “Quantum Brascamp–Lieb dualities”. In: *Communications in Mathematical Physics* (2023), pp. 1–24. arXiv: [1909.02383v3](https://arxiv.org/abs/1909.02383v3).
- [Bür+19] Peter Bürgisser et al. “Towards a theory of non-commutative optimization: Geodesic 1st and 2nd order methods for moment maps and polytopes”. In: *2019 IEEE 60th Annual Symposium on Foundations of Computer Science (FOCS)*. IEEE. 2019, pp. 845–861.
- [Bur69] Donald Bures. “An extension of Kakutani’s theorem on infinite product measures to the tensor product of semifinite  $w^*$ -algebras”. In: *Transactions of the American Mathematical Society* 135 (1969), pp. 199–212.
- [Buž+98] V Bužek et al. “Reconstruction of quantum states of spin systems: From quantum Bayesian inference to quantum tomography”. In: *Annals of Physics* 266.2 (1998), pp. 454–496.
- [BŻ17] Ingemar Bengtsson and Karol Życzkowski. *Geometry of quantum states: an introduction to quantum entanglement*. Cambridge university press, 2017.
- [Cad+14] Josh Cadney et al. “Inequalities for the ranks of multipartite quantum states”. In: *Linear Algebra and its Applications* 452 (2014), pp. 153–171.
- [Cas+21] Federico Castillo et al. *An effective solution to convex 1-body  $N$ -representability*. 2021. arXiv: [2105.06459v2](https://arxiv.org/abs/2105.06459v2).
- [CDW23] Matthias Christandl, Bergfinnur Durhuus, and Lasse Harboe Wolff. “The Quantum Entropy Cone near its Apex”. In: (2023). arXiv: [2306.00199](https://arxiv.org/abs/2306.00199).
- [CFS02] Carlton M Caves, Christopher A Fuchs, and Rüdiger Schack. “Unknown quantum states: the quantum de Finetti representation”. In: *Journal of Mathematical Physics* 43.9 (2002), pp. 4537–4559.
- [Che+14] Jianxin Chen et al. “Symmetric extension of two-qubit states”. In: *Physical Review A* 90.3 (2014), p. 032318.
- [Che+16] Jianxin Chen et al. “Detecting consistency of overlapping quantum marginals by separability”. In: *Physical Review A* 93.3 (2016), p. 032105. arXiv: [1509.06591v2](https://arxiv.org/abs/1509.06591v2).
- [Chi+04] Giulio Chiribella et al. “Covariant quantum measurements that maximize the likelihood”. In: *Physical Review A* 70.6 (2004), p. 062105.
- [Chi10] Giulio Chiribella. “On quantum estimation, quantum cloning and finite quantum de Finetti theorems”. In: *Conference on Quantum Computation, Communication, and Cryptography*. Springer. 2010, pp. 9–25. arXiv: [1010.1875v4](https://arxiv.org/abs/1010.1875v4).
- [CHM07] Matthias Christandl, Aram W. Harrow, and Graeme Mitchison. “Nonzero Kronecker Coefficients and What They Tell us about Spectra”. In: *Communications in Mathematical Physics* 270.3 (Jan. 2007), pp. 575–585.
- [Chr+07] Matthias Christandl et al. “One-and-a-half quantum de Finetti theorems”. In: *Communications in Mathematical Physics* 273.2 (2007), pp. 473–498.

- [Chr+14] Matthias Christandl et al. “Eigenvalue distributions of reduced density matrices”. In: *Communications in mathematical physics* 332.1 (2014), pp. 1–52. arXiv: [1204.0741v2](#).
- [CK10] Bob Coecke and Aleks Kissinger. “The compositional structure of multipartite quantum entanglement”. In: *International Colloquium on Automata, Languages, and Programming*. Springer. 2010, pp. 297–308. arXiv: [1002.2540v2](#).
- [CKR09] Matthias Christandl, Robert König, and Renato Renner. “Postselection technique for quantum channels with applications to quantum cryptography”. In: *Physical review letters* 102.2 (2009), p. 020504.
- [CKW00] Valerie Coffman, Joydip Kundu, and William K Wootters. “Distributed entanglement”. In: *Physical Review A* 61.5 (2000), p. 052306.
- [CLL13] Eric A Carlen, Joel L Lebowitz, and Elliott H Lieb. “On an extension problem for density matrices”. In: *Journal of Mathematical Physics* 54.6 (2013), p. 062103. arXiv: [1301.4605v1](#).
- [CLO13] David Cox, John Little, and Donal OShea. *Ideals, varieties, and algorithms: an introduction to computational algebraic geometry and commutative algebra*. Springer Science & Business Media, 2013.
- [CLR88] Wojciech Cegła, JT Lewis, and GA Raggio. “The free energy of quantum spin systems and large deviations”. In: *Communications in mathematical physics* 118.2 (1988), pp. 337–354.
- [CM06] Matthias Christandl and Graeme Mitchison. “The spectra of quantum states and the Kronecker coefficients of the symmetric group”. In: *Communications in mathematical physics* 261.3 (2006), pp. 789–797.
- [CM21] Benoît Collins and Colin McSwiggen. *Projections of Orbital Measures and Quantum Marginal Problems*. 2021. arXiv: [2112.13908](#).
- [Col01] A John Coleman. “Reduced density matrices—then and now”. In: *International Journal of Quantum Chemistry* 85.4-5 (2001), pp. 196–203.
- [Col63] A John Coleman. “Structure of fermion density matrices”. In: *Reviews of modern Physics* 35.3 (1963), p. 668.
- [Cou60] Charles A Coulson. “Present state of molecular structure calculations”. In: *Reviews of Modern Physics* 32.2 (1960), p. 170.
- [CR12] Matthias Christandl and Renato Renner. “Reliable quantum state tomography”. In: *Physical Review Letters* 109.12 (2012), p. 120403. arXiv: [1108.5329v2](#).
- [CŞW18] Matthias Christandl, M Burak Şahinoğlu, and Michael Walter. “Recoupling coefficients and quantum entropies”. In: *Annales Henri Poincaré*. Vol. 19. 2. Springer. 2018, pp. 385–410.
- [Cvi08] Predrag Cvitanović. *Group theory: birdtracks, Lie’s, and exceptional groups*. Princeton University Press, 2008.

- [CVZ18] Matthias Christandl, Péter Vrana, and Jeroen Zuiddam. “Universal points in the asymptotic spectrum of tensors”. In: *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing*. 2018, pp. 289–296.
- [CW20] Jordan Cotler and Frank Wilczek. “Quantum overlapping tomography”. In: *Physical review letters* 124.10 (2020), p. 100401.
- [CY00] AJ Coleman and VI Yukalov. “Reduced Density Matrices: Coulson’s Challenge”. In: *Many-Electron Densities and Reduced Density Matrices* (2000).
- [CY02] Terence H Chan and Raymond W Yeung. “On a relation between information inequalities and group theory”. In: *IEEE Transactions on Information Theory* 48.7 (2002), pp. 1992–1995.
- [DBE98] Radoslav Derka, Vladimir Buzek, and Artur K Ekert. “Universal algorithm for optimal estimation of quantum states from finite ensembles via realizable generalized measurement”. In: *Physical Review Letters* 80.8 (1998), p. 1571.
- [DE11] Paul Dupuis and Richard S Ellis. *A weak convergence approach to the theory of large deviations*. Vol. 902. John Wiley & Sons, 2011.
- [Den08] Frank Den Hollander. *Large deviations*. Vol. 14. American Mathematical Soc., 2008.
- [DH05] Sumit Daftuar and Patrick Hayden. “Quantum state transformations and the Schubert calculus”. In: *Annals of Physics* 315.1 (2005), pp. 80–122.
- [DLN20] Stephane Dartois, Luca Lioni, and Ion Nechita. “The joint distribution of the marginals of multipartite random quantum states”. In: *Random Matrices: Theory and Applications* 9.03 (2020), p. 2050010. arXiv: [1808.08554v2](https://arxiv.org/abs/1808.08554v2).
- [Doh+08] Andrew C Doherty et al. “The quantum moment problem and bounds on entangled multi-prover games”. In: *2008 23rd Annual IEEE Conference on Computational Complexity*. IEEE. 2008, pp. 199–210. arXiv: [0803.4373v1](https://arxiv.org/abs/0803.4373v1).
- [DPS03] G Mauro D’Ariano, Matteo GA Paris, and Massimiliano F Sacchi. “Quantum tomography”. In: *Advances in Imaging and Electron Physics* 128 (2003), pp. 206–309.
- [DPS04] Andrew C Doherty, Pablo A Parrilo, and Federico M Spedalieri. “Complete family of separability criteria”. In: *Physical Review A* 69.2 (2004), p. 022308.
- [Duf90] Nicholas G Duffield. “A large deviation principle for the reduction of product representations”. In: *Proceedings of the American Mathematical Society* 109.2 (1990), pp. 503–515.
- [DZ10] Amir Dembo and Ofer Zeitouni. *Large deviations techniques and applications*. Berlin: Springer, 2010.
- [Eis+08] Jens Eisert et al. “Gaussian quantum marginal problem”. In: *Communications in mathematical physics* 280.1 (2008), pp. 263–280.
- [Fan57] Ugo Fano. “Description of states in quantum mechanics by density matrix and operator techniques”. In: *Reviews of modern physics* 29.1 (1957), p. 74.



- [FC12] Tobias Fritz and Rafael Chaves. “Entropic inequalities and marginal problems”. In: *IEEE transactions on information theory* 59.2 (2012), pp. 803–817. arXiv: [1112.4788v2](#).
- [FGP21] Tobias Fritz, Tomáš Gonda, and Paolo Perrone. *De Finetti’s Theorem in Categorical Probability*. 2021. arXiv: [2105.02639](#).
- [FH13] William Fulton and Joe Harris. *Representation theory: a first course*. Vol. 129. Springer Science & Business Media, 2013.
- [Fin02] William Finnoff. “Integration of large-deviation kernels and applications to large deviations for evolutionary games”. In: *Probability theory and related fields* 122.1 (2002), pp. 141–162.
- [Fis05] Steve Fisk. “A very short proof of Cauchy’s interlace theorem for eigenvalues of Hermitian matrices”. In: (2005). arXiv: [math/0502408](#).
- [Fra22] Thomas C Fraser. *A sufficient family of necessary inequalities for the compatibility of quantum marginals*. 2022. arXiv: [2211.00685v1](#).
- [Ful00] William Fulton. “Eigenvalues, invariant factors, highest weights, and Schubert calculus”. In: *Bulletin of the American Mathematical Society* 37.3 (2000), pp. 209–249.
- [FW18] Thomas C Fraser and Elie Wolfe. “Causal compatibility inequalities admitting quantum violations in the triangle structure”. In: *Physical Review A* 98.2 (2018), p. 022113. arXiv: [1709.06242v2](#).
- [FW20] Cole Franks and Michael Walter. *Minimal length in an orbit closure as a semiclassical limit*. 2020. arXiv: [2004.14872v2](#).
- [Gao15] Jingliang Gao. “Quantum union bounds for sequential projective measurements”. In: *Physical Review A* 92.5 (2015), p. 052331. arXiv: [1410.5688v2](#).
- [Gar+17] Ankit Garg et al. “Algorithmic and optimization aspects of Brascamp-Lieb inequalities, via operator scaling”. In: *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing*. 2017, pp. 397–409. arXiv: [1607.06711v4](#).
- [GM05] Richard D Gill and Serge Massar. “State estimation for large ensembles”. In: *Asymptotic Theory of Quantum Statistical Inference: Selected Papers*. World Scientific, 2005, pp. 178–214.
- [Gur04] Leonid Gurvits. “Classical complexity and quantum entanglement”. In: *Journal of Computer and System Sciences* 69.3 (2004), pp. 448–484.
- [GW10] Gilad Gour and Nolan R Wallach. “All maximally entangled four-qubit states”. In: *Journal of Mathematical Physics* 51.11 (2010).
- [GŻ13] Eugene Gutkin and Karol Życzkowski. “Joint numerical ranges, quantum maps, and joint numerical shadows”. In: *Linear Algebra and its Applications* 438.5 (2013), pp. 2394–2404.

- [Haa+16] Jeongwan Haah et al. “Sample-optimal tomography of quantum states”. In: *Proceedings of the forty-eighth annual ACM symposium on Theory of Computing*. 2016, pp. 913–925. arXiv: [1508.01797v2](#).
- [Haa+21] Erkkka Haapasalo et al. “Quantum marginal problem and incompatibility”. In: *Quantum* 5 (2021), p. 476. arXiv: [1909.02941v3](#).
- [hai33] Alfred hair. “The notion of measure in the theory of continuous groups”. In: *Annals of mathematics* (1933), pp. 147–169.
- [Hal07] William Hall. “Compatibility of subsystem states and convex geometry”. In: *Physical Review A* 75.3 (2007), p. 032102.
- [Hal15] Brian C. Hall. *Lie Groups, Lie Algebras, and Representations: An Elementary Introduction*. 2nd ed. Graduate Texts in Mathematics 222. Springer International Publishing, 2015.
- [Har13] Aram W Harrow. *The church of the symmetric subspace*. 2013. arXiv: [1308.6595v1](#).
- [Har64] JOHN Hartigan. “Invariant prior distributions”. In: *The Annals of Mathematical Statistics* (1964), pp. 836–845.
- [Hay+04] Patrick Hayden et al. “Structure of states which satisfy strong subadditivity of quantum entropy with equality”. In: *Communications in mathematical physics* 246.2 (2004), pp. 359–374.
- [Hay01] Masahito Hayashi. “Asymptotics of quantum relative entropy from a representation theoretical viewpoint”. In: *Journal of Physics A: Mathematical and General* 34.16 (2001), p. 3413.
- [Hay02] Masahito Hayashi. “Two quantum analogues of Fisher information from a large deviation viewpoint of quantum estimation”. In: *Journal of Physics A: Mathematical and General* 35.36 (2002), p. 7689.
- [Hay05] Masahito Hayashi. *Asymptotic theory of quantum statistical inference: selected papers*. World Scientific, 2005.
- [HC13] Wolfram Helwig and Wei Cui. “Absolutely maximally entangled states: existence and applications”. In: (2013). arXiv: [1306.2536](#).
- [Hec82] Gert J Heckman. “Projections of orbits and asymptotic behavior of multiplicities for compact connected Lie groups”. In: *Inventiones mathematicae* 67.2 (1982), pp. 333–356.
- [Hel69] Carl W Helstrom. “Quantum detection and estimation theory”. In: *Journal of Statistical Physics* 1 (1969), pp. 231–252.
- [HGS17] Felix Huber, Otfried Gühne, and Jens Siewert. “Absolutely maximally entangled states of seven qubits do not exist”. In: *Physical review letters* 118.20 (2017), p. 200502.
- [Hig03] Atsushi Higuchi. “On the one-particle reduced density matrices of a pure three-qutrit quantum state”. In: (2003). arXiv: [quant-ph/0309186](#).



- [HJ85] RA Horn and CR Johnson. “Matrix Analysis, Cambridge Univ”. In: *Press. MR0832183* (1985).
- [HM76] Robin L Hudson and Graham R Moody. “Locally normal symmetric states and an analogue of de Finetti’s theorem”. In: *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete* 33.4 (1976), pp. 343–351.
- [HMZ16] Teiko Heinosaari, Takayuki Miyadera, and Mário Ziman. “An invitation to quantum incompatibility”. In: *Journal of Physics A: Mathematical and Theoretical* 49.12 (2016), p. 123001.
- [Hol11] Alexander S Holevo. *Probabilistic and statistical aspects of quantum theory*. Vol. 1. Springer Science & Business Media, 2011.
- [Hol78] AS Holevo. “Estimation of shift parameters of a quantum state”. In: *Reports on Mathematical Physics* 13.3 (1978), pp. 379–399.
- [How87] Roger Howe. “(GL(n), GL(m))-duality and symmetric plethysm”. In: *Proceedings of the Indian Academy of Sciences-Mathematical Sciences*. Vol. 97. 1-3. Springer. 1987, pp. 85–109.
- [HP91] Fumio Hiai and Dénes Petz. “The proper formula for relative entropy and its asymptotics in quantum probability”. In: *Communications in mathematical physics* 143.1 (1991), pp. 99–114.
- [Hra+04] Zdeněk Hradil et al. “3 maximum-likelihood methods in quantum mechanics”. In: *Quantum state estimation*. Springer, 2004, pp. 59–112.
- [Hra97] Zdenek Hradil. “Quantum-state estimation”. In: *Physical Review A* 55.3 (1997), R1561.
- [HSS03] Atsushi Higuchi, Anthony Sudbery, and Jason Szulc. “One-qubit reduced states of a pure many-qubit state: polygon inequalities”. In: *Physical review letters* 90.10 (2003), p. 107902.
- [Hub17] Felix Michael Huber. “Quantum states and their marginals: from multipartite entanglement to quantum error-correcting codes”. In: (2017).
- [HZG05] Y-J Han, Y-S Zhang, and G-C Guo. “Compatibility relations between the two-party reduced and global tripartite density matrices”. In: *Physical Review A* 72.5 (2005), p. 054302.
- [Iwa49] Kenkichi Iwasawa. “On some types of topological groups”. In: *Annals of Mathematics* (1949), pp. 507–558.
- [Jef98] Harold Jeffreys. *The theory of probability*. OuP Oxford, 1998.
- [Jir95] Mats Jirstrand. *Cylindrical algebraic decomposition-an introduction*. Linköping University, 1995.
- [Jon91] K. R. W. Jones. “Principles of quantum inference”. In: *Annals of Physics* 207.1 (1991), pp. 140–170.
- [Jon96] Andrew R Jones. “A combinatorial approach to the double cosets of the symmetric group with respect to Young subgroups”. In: *European Journal of Combinatorics* 17.7 (1996), pp. 647–655.

- [Key06] Michael Keyl. “Quantum state estimation and large deviations”. In: *Reviews in Mathematical Physics* 18.01 (2006), pp. 19–60.
- [Kim20] Isaac H Kim. *Entropy scaling law and the quantum marginal problem*. 2020. arXiv: [2010.07424](#).
- [Kla17] Joel David Klassen. “Existence and Uniqueness in the Quantum Marginal Problem”. PhD thesis. 2017.
- [Kly02] Alexander Klyachko. “Coherent states, entanglement, and geometric invariant theory”. In: (2002). arXiv: [quant-ph/0206012](#).
- [Kly04] Alexander Klyachko. *Quantum marginal problem and representations of the symmetric group*. 2004. arXiv: [quant-ph/0409113](#).
- [Kly06] Alexander A Klyachko. “Quantum marginal problem and N-representability”. In: *Journal of Physics: Conference Series*. Vol. 36. 1. IOP Publishing. 2006, p. 72.
- [Kly09] Alexander A Klyachko. *The Pauli exclusion principle and beyond*. 2009. arXiv: [0904.2009v1](#).
- [KN79] George Kempf and Linda Ness. “The length of vectors in representation spaces”. In: *Algebraic Geometry: Summer Meeting, Copenhagen, August 7–12, 1978*. Springer. 1979, pp. 233–243.
- [Kna01] Anthony W Knapp. “Representation theory of semisimple groups: an overview based on examples”. In: (2001).
- [KR05] Robert König and Renato Renner. “A de Finetti representation for finite symmetric quantum states”. In: *Journal of Mathematical physics* 46.12 (2005), p. 122108.
- [KS06] Alexander A Klyachko and Alexander S Shumovsky. “General entanglement”. In: *Journal of Physics: Conference Series*. Vol. 36. 1. IOP Publishing. 2006, p. 015.
- [KS13] Christian Kleiber and Jordan Stoyanov. “Multivariate distributions and the moment problem”. In: *Journal of Multivariate Analysis* 113 (2013), pp. 7–18.
- [Kul97] Solomon Kullback. *Information theory and statistics*. Courier Corporation, 1997.
- [KW01] Michael Keyl and Reinhard F Werner. “Estimating the spectrum of a density operator”. In: *Physical Review A* 64.5 (2001), p. 052311.
- [LC07] Hua Li and Edwin KP Chong. *A Group Theoretic Model for Information*. 2007. arXiv: [0710.1254v1](#).
- [LCV07] Yi-Kai Liu, Matthias Christandl, and Frank Verstraete. “Quantum computational complexity of the N-representability problem: QMA complete”. In: *Physical review letters* 98.11 (2007), p. 110503.
- [Lee01] Dong Hoon Lee. *The structure of complex Lie groups*. CRC Press, 2001.

- [LG22] Laurens T Ligthart and David Gross. *The inflation hierarchy and the polarization hierarchy are complete for the quantum bilocal scenario*. 2022. arXiv: [2212.11299v2](#).
- [LGG21] Laurens T Ligthart, Mariami Gachechiladze, and David Gross. *A convergent inflation hierarchy for quantum causal structures*. 2021. arXiv: [2110.14659v3](#).
- [Liu06] Yi-Kai Liu. “Consistency of local density matrices is QMA-complete”. In: *Approximation, randomization, and combinatorial optimization. algorithms and techniques*. Springer, 2006, pp. 438–449.
- [LR34] Dudley Ernest Littlewood and Archibald Read Richardson. “Group characters and algebra”. In: *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character* 233.721-730 (1934), pp. 99–141.
- [LS87] James Lynch and Jayaram Sethuraman. “Large deviations for processes with independent increments”. In: *The annals of probability* 15.2 (1987), pp. 610–627.
- [LSW11] Yeong-Cherng Liang, Robert W Spekkens, and Howard M Wiseman. “Specker’s parable of the overprotective seer: A road to contextuality, nonlocality and complementarity”. In: *Physics Reports* 506.1-2 (2011), pp. 1–39.
- [LTC+13] Eduardo V Lude, F Javier Torres, Cesar Costa, et al. “Functional n-representability in 2-matrix, 1-matrix, and density functional theories”. In: (2013).
- [LW05] Noah Linden and Andreas Winter. “A new inequality for the von Neumann entropy”. In: *Communications in mathematical physics* 259.1 (2005), pp. 129–138.
- [LW17] Cécilia Lancien and Andreas Winter. “Flexible constrained de Finetti reductions and applications”. In: *Journal of Mathematical Physics* 58.9 (2017), p. 092203.
- [Maj18] Christian Majenz. *Constraints on multipartite quantum entropies*. 2018. arXiv: [1810.12845v1](#).
- [Mal88] Francesco M Malvestuto. “Existence of extensions and product extensions for discrete probability distributions”. In: *Discrete Mathematics* 69.1 (1988), pp. 61–77.
- [Mar12] Iman Marvian. “Symmetry, asymmetry and quantum information”. In: (2012).
- [Maz12a] David A Mazziotti. “Significant conditions for the two-electron reduced density matrix from the constructive solution of N representability”. In: *Physical Review A* 85.6 (2012), p. 062507. arXiv: [1207.0541v1](#).
- [Maz12b] David A Mazziotti. “Structure of fermionic density matrices: Complete N-representability conditions”. In: *Physical Review Letters* 108.26 (2012), p. 263002. arXiv: [1112.5866v2](#).

- [MFK94] David Mumford, John Fogarty, and Frances Kirwan. *Geometric invariant theory*. Vol. 34. Springer Science & Business Media, 1994.
- [Mil14] JS Milne. *Algebraic groups*. 2014.
- [Mit07] Graeme Mitchison. “A dual de Finetti theorem”. In: (2007). arXiv: [quant-ph/0701064](#).
- [MN84] David Bryant Mumford and Linda Ness. “A Stratification of the Null Cone Via the Moment Map”. In: *American Journal of Mathematics* (1984).
- [MP95] Serge Massar and Sandu Popescu. “Optimal extraction of information from finite quantum ensembles”. In: *Physical review letters* 74.8 (1995), p. 1259.
- [MW13] Ashley Montanaro and Ronald de Wolf. *A survey of quantum property testing*. 2013. arXiv: [1310.2035v4](#).
- [NBA21] Miguel Navascués, Flavio Baccari, and Antonio Acín. “Entanglement marginal problems”. In: *Quantum* 5 (2021), p. 589.
- [NPA08] Miguel Navascués, Stefano Pironio, and Antonio Acín. “A convergent hierarchy of semidefinite programs characterizing the set of quantum correlations”. In: *New Journal of Physics* 10.7 (2008), p. 073013.
- [Ock+11] Samuel A Ocko et al. “Quantum codes give counterexamples to the unique preimage conjecture of the N-representability problem”. In: *Physical Review Letters* 106.11 (2011), p. 110501. arXiv: [1010.2717v5](#).
- [ON05] Tomohiro Ogawa and Hiroshi Nagaoka. “Strong converse and Stein’s lemma in quantum hypothesis testing”. In: *Asymptotic Theory of Quantum Statistical Inference: Selected Papers*. World Scientific, 2005, pp. 28–42.
- [Osb08] Tobias J Osborne. *Entropic bounds for the quantum marginal problem*. 2008. arXiv: [0806.2962v1](#).
- [OV12] Arkadij L Onishchik and Ernest B Vinberg. *Lie groups and algebraic groups*. Springer Science & Business Media, 2012.
- [OW16] Ryan O’Donnell and John Wright. “Efficient quantum tomography”. In: *Proceedings of the forty-eighth annual ACM symposium on Theory of Computing*. 2016, pp. 899–912. arXiv: [1508.01907v2](#).
- [Pag93] Don N Page. “Average entropy of a subsystem”. In: *Physical review letters* 71.9 (1993), p. 1291.
- [PBP22] Jason L Pereira, Leonardo Banchi, and Stefano Pirandola. *Analytical bounds for non-asymptotic asymmetric state discrimination*. 2022. arXiv: [2207.10699v2](#).
- [Pen71] Roger Penrose. “Applications of negative dimensional tensors”. In: *Combinatorial mathematics and its applications* 1 (1971), pp. 221–244.
- [Per72] Askold M Perelomov. “Coherent states for arbitrary Lie group”. In: *Communications in Mathematical Physics* 26 (1972), pp. 222–236.
- [Pet03] Dénes Petz. “Monotonicity of quantum relative entropy revisited”. In: *Reviews in Mathematical Physics* 15.01 (2003), pp. 79–91.

- [Pip03] Nicholas Pippenger. “The inequalities of quantum information theory”. In: *IEEE Transactions on Information Theory* 49.4 (2003), pp. 773–789.
- [PNA10] Stefano Pironio, Miguel Navascués, and Antonio Acín. “Convergent relaxations of polynomial optimization problems with noncommuting variables”. In: *SIAM Journal on Optimization* 20.5 (2010), pp. 2157–2180.
- [PPY19] Igor Pak, Greta Panova, and Damir Yeliussizov. “On the largest Kronecker and Littlewood–Richardson coefficients”. In: *Journal of Combinatorial Theory, Series A* 165 (2019), pp. 44–77. arXiv: [1804.04693v2](#).
- [PR04] Matteo Paris and Jaroslav Rehacek. *Quantum state estimation*. Vol. 649. Springer Science & Business Media, 2004.
- [Pro07] Claudio Procesi. “Lie groups. An approach through invariants and representations”. In: *Bull. Amer. Math. Soc* (2007).
- [Ren10] Renato Renner. “Simplifying information-theoretic arguments by post-selection.” In: (2010).
- [Rob29] Howard Percy Robertson. “The uncertainty principle”. In: *Physical Review* 34.1 (1929), p. 163.
- [Rud53] Walter Rudin. *Principles of mathematical analysis*. 1953.
- [Rus07] Mary Beth Ruskai. “Connecting N-representability to Weyl’s problem: the one-particle density matrix for N= 3 and R= 6”. In: *Journal of Physics A: Mathematical and Theoretical* 40.45 (2007), F961. arXiv: [0706.1855v1](#).
- [RW09] Mark D Reid and Robert C Williamson. *Generalised pinsker inequalities*. 2009. arXiv: [0906.1244v1](#).
- [Ryb19] Christopher Ryba. *A Permutation Module Deligne Category and Stable Patterns of Kronecker Coefficients*. 2019. arXiv: [1909.04100v1](#).
- [Sag13] Bruce E Sagan. *The symmetric group: representations, combinatorial algorithms, and symmetric functions*. Vol. 203. Springer Science & Business Media, 2013.
- [San61] Ivan Nicolaevich Sanov. “On the probability of large deviations of random variables”. In: *Selected Translations in Mathematical Statistics and Probability* 1 (1961), pp. 213–244.
- [SBC01] Ruediger Schack, Todd A Brun, and Carlton M Caves. “Quantum bayes rule”. In: *Physical Review A* 64.1 (2001), p. 014305.
- [Sch15] Christian Schilling. *Quantum marginal problem and its physical relevance*. 2015. arXiv: [1507.00299](#).
- [Sco04] Andrew J Scott. “Multipartite entanglement, quantum-error-correcting codes, and entangling power of quantum evolutions”. In: *Physical Review A* 69.5 (2004), p. 052330.
- [Sel12] Peter Selinger. *Finite dimensional Hilbert spaces are complete for dagger compact closed categories*. 2012. arXiv: [1207.6972v2](#).

- [SGC13] Christian Schilling, David Gross, and Matthias Christandl. “Pinning of fermionic occupation numbers”. In: *Physical review letters* 110.4 (2013), p. 040404. arXiv: [1210.5531v1](#).
- [Sha48] Claude Elwood Shannon. “A mathematical theory of communication”. In: *The Bell system technical journal* 27.3 (1948), pp. 379–423.
- [Sim96] Barry Simon. *Representations of finite and compact groups*. 10. American Mathematical Soc., 1996.
- [Sra16] Suvrit Sra. “On inequalities for normalized Schur functions”. In: *European Journal of Combinatorics* 51 (2016), pp. 492–494. arXiv: [1502.04753v3](#).
- [Ste97] J Michael Steele. *Probability theory and combinatorial optimization*. SIAM, 1997.
- [Sto13] Jordan M Stoyanov. *Counterexamples in probability*. Courier Corporation, 2013.
- [Str16] Stanislav Sergeevich Straupe. “Adaptive quantum tomography”. In: *JETP letters* 104 (2016), pp. 510–522.
- [Tao11] Terence Tao. *An introduction to measure theory*. Vol. 126. American Mathematical Society Providence, 2011.
- [Tou11] Hugo Touchette. *A basic introduction to large deviations: Theory, applications, simulations*. 2011. arXiv: [1106.4146v3](#).
- [TV15] Tomáš Tyc and Jan Vlach. “Quantum marginal problems”. In: *The European Physical Journal D* 69.9 (2015), pp. 1–6.
- [TV99] Rolf Tarrach and Guifré Vidal. “Universality of optimal measurements”. In: *Physical Review A* 60.5 (1999), R3339.
- [Var66] SR Srinivasa Varadhan. “Asymptotic probabilities and differential equations”. In: *Communications on Pure and Applied Mathematics* 19.3 (1966), pp. 261–286.
- [VB96] Lieven Vandenberghe and Stephen Boyd. “Semidefinite programming”. In: *SIAM review* 38.1 (1996), pp. 49–95.
- [Ver+23] Benoît Vermersch et al. “Enhanced estimation of quantum properties with common randomized measurements”. In: (2023). arXiv: [2304.12292](#).
- [Vla15] Jan Vlach. “Quantum marginal problem”. PhD thesis. Masarykova univerzita, Fakulta informatiky, 2015.
- [Vor62] Nikolai Nikolaevich Vorob’ev. “Consistent families of measures and their extensions”. In: *Theory of Probability & Its Applications* 7.2 (1962), pp. 147–163.
- [VR89] K Vogel and H Risken. “Determination of quasiprobability distributions in terms of probability distributions for the rotated quadrature phase”. In: *Physical Review A* 40.5 (1989), p. 2847.



- [VZ13] L Campos Venuti and Paolo Zanardi. “Probability density of quantum expectation values”. In: *Physics Letters A* 377.31-33 (2013), pp. 1854–1861. arXiv: [1202.4810v3](#).
- [Wal+13] Michael Walter et al. “Entanglement Polytopes: Multipartite Entanglement from Single-Particle Information”. In: *Science* 340.6137 (2013), pp. 1205–1208. arXiv: [1208.0365v4](#).
- [Wal14] Michael Walter. *Multipartite quantum states and their marginals*. 2014. arXiv: [1410.6820v1](#).
- [Wal17] Nolan R Wallach. “Geometric invariant theory”. In: *Universitext. Cham: Springer* (2017).
- [WBC11] Christopher J Wood, Jacob D Biamonte, and David G Cory. *Tensor networks and graphical calculus for open quantum systems*. 2011. arXiv: [1111.6950v3](#).
- [Weh78] Alfred Wehrl. “General properties of entropy”. In: *Reviews of Modern Physics* 50.2 (1978), p. 221.
- [Wer18] Konstantin Wernli. *Computing Entanglement Polytopes*. 2018. arXiv: [1808.03382v1](#).
- [Wey46] Hermann Weyl. *The classical groups: their invariants and representations*. 1. Princeton university press, 1946.
- [Wey50] Hermann Weyl. *The theory of groups and quantum mechanics*. Trans. by H. P. Robertson. Courier Corporation, 1950.
- [WMN10] Tzu-Chieh Wei, Michele Mosca, and Ashwin Nayak. “Interacting boson problems can be QMA hard”. In: *Physical review letters* 104.4 (2010), p. 040501.
- [Woo10] Chris Woodward. “Moment maps and geometric invariant theory”. In: *Les cours du CIRM* 1.1 (2010), pp. 55–98. arXiv: [0912.1132v6](#).
- [Woo90] William K Wootters. “Random quantum states”. In: *Foundations of Physics* 20.11 (1990), pp. 1365–1378.
- [WWB17] Peter Woit, Woit, and Bartolini. *Quantum theory, groups and representations*. Springer, 2017.
- [WZ14] John Archibald Wheeler and Wojciech Hubert Zurek. *Quantum theory and measurement*. Vol. 40. Princeton University Press, 2014.
- [Yu+21] Xiao-Dong Yu et al. “A complete hierarchy for the pure state marginal problem in quantum mechanics”. In: *Nature communications* 12.1 (2021), pp. 1–7.
- [Zha+22] Lin Zhang et al. “Probability density functions of quantum mechanical observable uncertainties”. In: *Communications in Theoretical Physics* 74.7 (2022), p. 075102. arXiv: [2205.03193v1](#).
- [Zhe73] Dmitriï Petrovich Zhelobenko. *Compact Lie groups and their representations*. Vol. 40. American Mathematical Soc., 1973.
- [ZRM21] Andrew Zhao, Nicholas C Rubin, and Akimasa Miyake. “Fermionic partial tomography via classical shadows”. In: *Physical Review Letters* 127.11 (2021), p. 110504. arXiv: [2010.16094v3](#).

- [ZS01] Karol Zyczkowski and Hans-Jürgen Sommers. “[Induced measures in the space of mixed quantum states](#)”. In: *Journal of Physics A: Mathematical and General* 34.35 (2001), p. 7111.
- [Zyc+98] Karol Zyczkowski et al. “On the volume of the set of mixed entangled states”. In: (1998). arXiv: [quant-ph/9804024](#).



# Appendices

## A.1 Misc. Results

**Lemma A.1.1.** *Let  $X$  be a topological space and let  $f : X \rightarrow [-\infty, \infty]$  be a function. Let  $L_f(c) = \{x \in X \mid f(x) \leq c\}$  be a lower level set of  $f$ . Then  $f$  is lower-semicontinuous if and only if  $L_f(c)$  is closed for all  $c$ .*

*Proof.* To prove the “only if” portion of the proof, note that a function  $f : X \rightarrow [-\infty, \infty]$  is lower-semicontinuous if and only if whenever  $f(x) > y$ , there exists an open neighborhood  $N \subseteq X$  of  $x$  such that  $\forall n \in N : f(n) > y$ . Since  $x \notin L_f(c)$  if and only if  $f(x) > c$  we conclude that there exists an open neighborhood  $N$  of  $x$  such that  $\forall n \in N : f(n) > c$  and thus  $N \subseteq X \setminus L_f(c)$ . Therefore, every  $x \in X \setminus L_f(c)$  is an interior point of  $X \setminus L_f(c)$  and thus  $L_f(c)$  is closed. To prove the “if” portion of the proof, note that if  $X \setminus L_f(c)$  is open for all  $c$  then for every  $x \in X \setminus L_f(c)$  there exists an open neighborhood  $N$  of  $x$  contained in  $X \setminus L_f(c)$  such that  $\forall n \in N, f(n) > c$  and thus  $f$  is lower-semicontinuous.  $\square$

## A.2 Invariant post-selection

At the core of many of the theorems and results of this thesis is a deceptively simple yet powerful technique for placing universal upper bounds on the probability,  $\text{Tr}(\rho E)$ , where the effect  $E \in \mathcal{B}(\mathcal{V})$  is  $K$ -invariant ( $k \cdot E = \Phi(k)^* E \Phi(k) = E$  for all  $k \in K$ ) for generic states  $\rho$  in terms of the probability assigned to  $E$  by a special state  $\tau$ . As will be shown in this section, this powerful technique relies upon viewing the state  $\rho$  as the post-measurement state obtained after witnessing a particular outcome when measuring a purification of the special state  $\tau$ . It is from perspective that this this technique was termed the *post-selection technique* by Christandl, König, and Renner when it was first introduced in [CKR09] and [Ren10]. Although the post-selection technique was only explicitly demonstrated in the context where  $G$  was the symmetric group  $S_n$  and  $\mathcal{V}$  was the  $n$ -fold tensor product space, the original authors accurately concluded that their concept could be extended to the more general setting presented here.

We begin by considering the following steering lemma which expresses every quantum state,  $\rho$ , as the resulting steered state after obtaining a outcome dual to  $\rho$  when measuring a portion of a maximally entangled state.

**Proposition A.2.1.** *Let  $\mathcal{H}$  be a finite-dimensional complex Hilbert space with dimension  $d = \dim(\mathcal{H})$  and let  $\mathcal{Z} \simeq \mathbb{C}^d$  be another finite-dimensional complex Hilbert space isomorphic to  $\mathcal{H}$*

through the identification of an orthonormal basis  $\{|1\rangle_{\mathcal{H}}, \dots, |d\rangle_{\mathcal{H}}\}$  for  $\mathcal{H}$  with an orthonormal basis  $\{|1\rangle_{\mathcal{Z}}, \dots, |d\rangle_{\mathcal{Z}}\}$  for  $\mathcal{Z}$ . Then any quantum state  $\rho_{\mathcal{H}} \in \mathcal{S}(\mathcal{H})$  on  $\mathcal{H}$  can be expressed as

$$\rho_{\mathcal{H}} = d \text{Tr}_{\mathcal{Z}}[(\mathbb{1}_{\mathcal{H}} \otimes D_{\mathcal{Z}}^{\rho})(|\psi\rangle\langle\psi|)_{\mathcal{H} \otimes \mathcal{Z}}] \quad (\text{A.2.1})$$

where  $|\psi\rangle_{\mathcal{H} \otimes \mathcal{Z}}$  is the pure state

$$|\psi\rangle_{\mathcal{H} \otimes \mathcal{Z}} = \frac{1}{\sqrt{d}} \sum_{i=1}^d |i\rangle_{\mathcal{H}} \otimes |i\rangle_{\mathcal{Z}}, \quad (\text{A.2.2})$$

and  $D_{\mathcal{Z}}^{\rho} \in \mathcal{B}(\mathcal{Z})$  is the effect dual to  $\rho$  defined by

$$D_{\mathcal{Z}}^{\rho} = \sum_{i,j=1}^d (|i\rangle\langle j|)_{\mathcal{Z}} \text{Tr}[\rho_{\mathcal{H}}(|j\rangle\langle i|)_{\mathcal{H}}]. \quad (\text{A.2.3})$$

The next result essentially extends [Proposition A.2.1](#) to the setting where  $\rho$  is a  $K$ -covariant quantum state. It can be seen as a generalization of [\[CKR09, Lem. 2\]](#) or equivalently [\[Ren10, Lem. 3\]](#).

**Corollary A.2.2.** *Let  $\Phi : K \rightarrow \text{U}(\mathcal{V})$  be a unitary representation of a compact group  $K$  on a finite-dimensional complex Hilbert space  $\mathcal{V}$ . Let  $\mathcal{V}^*$  denote the dual vector space of  $\mathcal{V}$  and let  $\Phi^* : K \rightarrow \text{U}(\mathcal{V}^*)$  be the dual representation of  $K$  on  $\mathcal{V}^*$ . Let  $\rho_{\mathcal{V}} \in \mathcal{S}(\mathcal{V})$  be a  $K$ -covariant quantum state in the sense that for all  $k \in K$ ,  $\Phi(k)\rho_{\mathcal{V}}\Phi(k^{-1}) = \rho_{\mathcal{V}}$ . Then there exists an ancillary Hilbert space  $\mathcal{Z}$ , state  $\tau_{\mathcal{V} \otimes \mathcal{Z}} \in \mathcal{S}(\mathcal{V} \otimes \mathcal{Z})$  and effect  $D_{\mathcal{Z}}^{\rho} \in \mathcal{B}(\mathcal{Z})$  satisfying*

$$\rho_{\mathcal{V}} = \dim((\mathcal{V} \otimes \mathcal{V}^*)^K) \text{Tr}_{\mathcal{Z}}[(\mathbb{1}_{\mathcal{V}} \otimes D_{\mathcal{Z}}^{\rho})\tau_{\mathcal{V} \otimes \mathcal{Z}}]. \quad (\text{A.2.4})$$

*Proof.* Consider the subspace of  $K$ -invariant vectors in  $\mathcal{V} \otimes \mathcal{V}^*$  defined by

$$(\mathcal{V} \otimes \mathcal{V}^*)^K := \{w \in \mathcal{V} \otimes \mathcal{V}^* \mid \forall k \in K : (\Phi(k) \otimes \Phi^*(k))w = w\}, \quad (\text{A.2.5})$$

and let its dimension be denoted by

$$\gamma := \dim((\mathcal{V} \otimes \mathcal{V}^*)^K). \quad (\text{A.2.6})$$

Let  $\tau_{\mathcal{V} \otimes \mathcal{V}^*} \in \mathcal{S}(\mathcal{V} \otimes \mathcal{V}^*)$  be the quantum state proportional to the projection operator  $\Pi_{\mathcal{V} \otimes \mathcal{V}^*}^K$  onto the fixed subspace  $(\mathcal{V} \otimes \mathcal{V}^*)^K$ :

$$\tau_{\mathcal{V} \otimes \mathcal{V}^*} = \gamma^{-1} \Pi_{\mathcal{V} \otimes \mathcal{V}^*}^K. \quad (\text{A.2.7})$$

Additionally, let  $\tau_{\mathcal{V}} \in \mathcal{S}(\mathcal{V})$  be partial trace (over  $\mathcal{V}^*$ ) of  $\tau_{\mathcal{V} \otimes \mathcal{V}^*}$ :

$$\tau_{\mathcal{V}} = \text{Tr}_{\mathcal{V}^*}[\tau_{\mathcal{V} \otimes \mathcal{V}^*}] = \gamma^{-1} \text{Tr}_{\mathcal{V}^*}[\Pi_{\mathcal{V} \otimes \mathcal{V}^*}^K]. \quad (\text{A.2.8})$$

Now let  $\{|1\rangle_{(\mathcal{V} \otimes \mathcal{V}^*)^K}, \dots, |\gamma\rangle_{(\mathcal{V} \otimes \mathcal{V}^*)^K}\}$  be an arbitrary orthonormal basis for the subspace  $(\mathcal{V} \otimes \mathcal{V}^*)^K \subset \mathcal{V} \otimes \mathcal{V}^*$  and fix a purifying space  $\mathcal{Z} \simeq \mathbb{C}^{\gamma}$  isomorphic to  $(\mathcal{V} \otimes \mathcal{V}^*)^K$  and consider the unit vector

$$|\psi\rangle_{\mathcal{V} \otimes \mathcal{V}^* \otimes \mathcal{Z}} := \frac{1}{\sqrt{\gamma}} \sum_{i=1}^{\gamma} |i\rangle_{(\mathcal{V} \otimes \mathcal{V}^*)^K} \otimes |i\rangle_{\mathcal{Z}}. \quad (\text{A.2.9})$$

Viewing  $|\psi\rangle_{\mathcal{V}\otimes\mathcal{V}^*\otimes\mathcal{Z}}$  as a unit vector in  $\mathcal{V}\otimes\mathcal{V}^*\otimes\mathcal{Z}$ , it becomes a purification of  $\tau_{\mathcal{V}\otimes\mathcal{V}^*}$  and thus its sensible to define

$$\tau_{\mathcal{V}\otimes\mathcal{V}^*\otimes\mathcal{Z}} := (|\psi\rangle\langle\psi|)_{\mathcal{V}\otimes\mathcal{V}^*\otimes\mathcal{Z}}. \quad (\text{A.2.10})$$

Finally let

$$\tau_{\mathcal{V}\otimes\mathcal{Z}} := \text{Tr}_{\mathcal{V}^*}[\tau_{\mathcal{V}\otimes\mathcal{V}^*\otimes\mathcal{Z}}]. \quad (\text{A.2.11})$$

Then for any  $\rho_{\mathcal{V}} \in \text{End}(\mathcal{V})^K \cong (\mathcal{V}\otimes\mathcal{V}^*)^K$ , the effect  $D_{\mathcal{Z}}^\rho \in \mathcal{B}(\mathcal{Z})$  considered in [Proposition A.2.1](#) (where  $\mathcal{H}$  is replaced by  $(\mathcal{V}\otimes\mathcal{V}^*)^K$ ) satisfies

$$\rho_{\mathcal{V}} = \gamma \text{Tr}_{\mathcal{Z}}[(\mathbb{1}_{\mathcal{V}} \otimes D_{\mathcal{Z}}^\rho) \tau_{\mathcal{V}\otimes\mathcal{Z}}], \quad (\text{A.2.12})$$

which proves the claim.  $\square$

**Theorem A.2.3.** *Let  $\Phi : K \rightarrow \text{U}(\mathcal{V})$  be a unitary representation of a compact group on a finite-dimensional complex Hilbert space  $\mathcal{V}$ . Let  $E \in \mathcal{B}(\mathcal{V})$  be an effect,  $0_{\mathcal{V}} \leq E \leq \mathbb{1}_{\mathcal{V}}$ , that is  $K$ -covariant in the sense that*

$$\forall k \in K \quad : \quad \Phi(k)E\Phi(k^{-1}) = E. \quad (\text{A.2.13})$$

Then there exists a state  $\tau$  such that

$$\sup_{\rho \in \mathcal{S}(\mathcal{V})} \text{Tr}[E\rho] \leq \dim((\mathcal{V}\otimes\mathcal{V}^*)^K) \text{Tr}[E\tau], \quad (\text{A.2.14})$$

where  $(\mathcal{V}\otimes\mathcal{V}^*)^K$  is the subspace of  $K$ -invariant vectors in  $\mathcal{V}\otimes\mathcal{V}^*$ .

*Proof.* By the  $K$ -covariance of  $E$ , for any state  $\rho \in \mathcal{S}(\mathcal{V})$ , the value of the probability  $\text{Tr}(E\rho)$  is always equal to the probability associated to a  $K$ -covariant state  $\tilde{\rho}$  because

$$\begin{aligned} \text{Tr}[E\rho] &= \int_K d\mu(k) \text{Tr}[\Phi(k)E\Phi(k^{-1})\rho], \\ &= \text{Tr}[E \int_G d\mu(k) \Phi(k^{-1})\rho\Phi(k)], \\ &= \text{Tr}[E\tilde{\rho}]. \end{aligned} \quad (\text{A.2.15})$$

Therefore, by an application of [Corollary A.2.2](#), there exists a state  $\tau_{\mathcal{V}\otimes\mathcal{Z}}$  such that

$$\text{Tr}[E\tilde{\rho}] = \gamma \text{Tr}[(E \otimes D_{\mathcal{Z}}^{\tilde{\rho}}) \tau_{\mathcal{V}\otimes\mathcal{Z}}] \quad (\text{A.2.16})$$

where  $\gamma = \dim((\mathcal{V}\otimes\mathcal{V}^*)^K)$ . The upper-bound on  $\text{Tr}[E\tilde{\rho}]$  then emerges by noting that  $D_{\mathcal{Z}}^{\tilde{\rho}} \leq \mathbb{1}_{\mathcal{Z}}$  and thus

$$\text{Tr}[E\tilde{\rho}] \leq \gamma \text{Tr}[(E \otimes \mathbb{1}_{\mathcal{Z}}) \tau_{\mathcal{V}\otimes\mathcal{Z}}] = \gamma \text{Tr}[E\tau_{\mathcal{V}}]. \quad (\text{A.2.17})$$

$\square$

### A.3 Expected expectation values

The purpose of this section is to calculate the probability density of the probability measure obtained by pushing forward the uniform probability measure over the unit sphere,  $S^{2d-1}$ , of a  $d$ -dimensional complex Hilbert space through the natural map into the projective space  $\mathbb{P}(\mathbb{C}^d)$  and then through the expectation value map  $\psi \mapsto \text{Tr}(P_\psi X)$  of a single observable  $X \in \text{End}(\mathcal{H})$  satisfying  $X^* = X$ .

It is important to note that the results obtained in this section, namely [Theorem A.3.1](#) and [Theorem A.3.2](#), are not new. These results were obtained previously by Venuti and Zanardi [[VZ13](#), Eq. (17)] and later by Zhang et al. [[Zha+22](#), Prop. 4]. These results can also be interpreted as a special case of [[Chr+14](#), Thm. 4.1]. The particular Laplace transform techniques used to derive this probability density are due to Zhang et al. [[Zha+22](#)].

Our reasoning for including the forthcoming derivation in this thesis is three-fold. The first reason is simply that the exact formula obtained here serves as a point of comparison to the asymptotic approximations obtained in [Section 5.1.3](#) and [Section 5.1.4](#). The second reason is simply to highlight, from the perspective of residue theory, the significant role played by the degeneracies in the spectrum of the observable  $X$ . Finally, the third reason is to expose some of the challenges that arise when trying to calculate the *joint* probability density associated to the expectation values of a tuple of non-commuting observables (see [Remark A.3.3](#)).

The starting point is to consider the natural  $U(d)$ -invariant probability measure over the sphere of unit vectors  $S^{2d-1} \simeq \{v \in \mathbb{C}^d \mid \|v\|_2 = 1\}$  ([Definition 2.2.6](#)), and its push-forward to a measure over the projective space  $\mathbb{P}(\mathbb{C}^d) \simeq \mathbb{C}P^{d-1}$  via the map  $v \mapsto [v] = \mathbb{C}v$ . The surface area of the unit sphere  $S^{2d-1}$  is known to be equal to  $2\pi^d/\Gamma(d)$  where  $\Gamma$  is the *Gamma function* which, when restricted to positive integers, has the form  $\Gamma(d) = (d-1)!$ . Therefore, the uniform measure on the unit sphere  $S^{2d-1} \subset \mathbb{R}^{2d}$  can be expressed as a probability density relative to the Lebesgue measure on  $\mathbb{R}^{2d}$  by

$$d\nu(v) = \frac{\Gamma(d)}{2\pi^d} \delta(1 - \|v\|_2) \prod_{i=1}^d dx_i dy_i, \quad (\text{A.3.1})$$

where  $\delta(1 - \|v\|_2)$  enforces normalization and where the coordinates  $x_i$  and  $y_i$  are related to  $v$  via  $z_j = x_j + iy_j = \langle e_j, v \rangle$  for some chosen orthonormal basis  $\{e_1, \dots, e_d\}$  of  $\mathbb{C}^d$ . Furthermore note that  $\delta(1 - \|v\|_2) = 2\delta(1 - \|v\|_2^2)$ .

The pushforward of the measure  $\nu$  on the unit sphere  $S^{2d-1}$  through the aforementioned map  $v \mapsto \mathbb{C}v$  yields a measure on  $\mathbb{P}(\mathbb{C}^d)$ , denoted by  $\mu$ , and can be defined implicitly for all measurable functions  $g : \mathbb{P}(\mathbb{C}^d) \rightarrow \mathbb{R}$  by the equation

$$\int_{\psi \in \mathbb{P}(\mathbb{C}^d)} d\mu(\psi) g(\psi) = \int_{v \in S^{2d-1}} d\nu(v) g([v]). \quad (\text{A.3.2})$$

Therefore if  $X$  is a Hermitian operator on  $\mathcal{H} \simeq \mathbb{C}^d$ , then the pushforward of  $\mu$  through the expectation value map  $\psi \mapsto \Gamma_\psi(X)$  has density  $f_X(x)$  defined by

$$f_X(x) = \int_{\psi \in \mathbb{P}(\mathbb{C}^d)} \delta(x - \Gamma_\psi(X)) d\mu(\psi) = \int_{v \in S^{2d-1}} \delta(x - \langle v, Xv \rangle) d\nu(v). \quad (\text{A.3.3})$$

Next let the eigenvalues of  $X$  be denoted by  $\{\lambda_1, \dots, \lambda_d\} \subset \mathbb{R}$  and then non-uniquely construct an orthonormal basis  $\{e_1, \dots, e_d\}$  for  $\mathbb{C}^d$  consisting of eigenvectors of  $X$  such that  $Xe_j = \lambda_j e_j$  holds for all  $j \in [d]$ . Using this basis in Eq. (A.3.1) with coordinates  $z_j = x_j + iy_j = r_j e^{i\theta_j} = \langle e_j, v \rangle$  simplifies the expression  $\langle v, Xv \rangle$  above substantially. In particular, by switching to polar coordinates (with change of measure  $dx_j dy_j = r_j dr_j d\theta_j$ ), then integrating over the polar angles  $\theta_j$ , and then finally changing variables to the non-negative quantity  $p_j = r_j^2 = |\langle e_j, v \rangle|^2$  produces

$$f_X(x) = \Gamma(d) \int_{p \in \mathbb{R}_{\geq 0}^d} \delta(x - \sum_{i=1}^d \lambda_i p_i) \delta(1 - \sum_{i=1}^d p_i) \prod_{i=1}^d dp_i. \quad (\text{A.3.4})$$

This integral expression for  $f_X(x)$  admits of a straightforward interpretation as an integral over the standard probability simplex, i.e., where the coordinates  $(p_1, \dots, p_d)$  form a probability distribution and  $\sum_{i=1}^d \lambda_i p_i = x$  expresses the constraint that  $x \in \mathbb{R}$  must be a convex combination of the spectra of  $X$ .

From here, one can directly compute the density  $f_X(x)$  by calculating its characteristic function via Fourier methods and then using Gaussian integration as in [VZ13], or equivalently by calculating its moment generating function via Laplace methods and then using the inverse Laplace transform as in [Zha+22]. As previously mentioned, we will adopt the Laplace method used by Zhang et al. [Zha+22].

Recall that the (bilateral) Laplace transform and inverse Laplace transforms (also known as the Bromwich or Fourier-Mellin integration) can be respectively written as

$$F(s) = \mathcal{L}\{f\}(s) = \int_{-\infty}^{\infty} f(t) e^{-st} dt, \quad (\text{A.3.5})$$

$$f(t) = \mathcal{L}^{-1}\{F\}(t) = \frac{1}{2\pi i} \lim_{T \rightarrow \infty} \int_{\sigma - iT}^{\sigma + iT} F(s) e^{st} ds. \quad (\text{A.3.6})$$

Note that the latter integral expression for  $\mathcal{L}^{-1}\{F\}(t)$  is calculated by means of residue theory. Specifically, the integral may be evaluated along any line segment from  $\sigma - iT$  to  $\sigma + iT$  in the complex plane with constant real part  $\sigma$  such that the integral expression for Laplace transform evaluated at  $\sigma$ , namely  $F(\sigma)$  above, actually converges. For our purposes, this condition amounts to ensuring that the value of  $\sigma$  is larger than all of the real parts of all poles of  $F(s)$  and such that  $F(s)$  is bounded along this line segment.

Returning to Eq. (A.3.4), the Laplace transform for  $f_X(x)$  can be interpreted as a moment generating function (with opposite sign) for the random variable with density  $f_X(x)$ . Specifically, the Laplace transform of  $f_X(x)$  as a function of  $\chi$  is

$$\mathcal{L}\{f_X\}(\chi) = \Gamma(d) \int_{p \in \mathbb{R}_{\geq 0}^d} \exp(-\chi \sum_{i=1}^d \lambda_i p_i) \delta(1 - \sum_{i=1}^d p_i) \prod_{i=1}^d dp_i. \quad (\text{A.3.7})$$

Following [Zha+22], we renormalize the  $p_i$ -coordinates such that  $\sum_i p_i = t$  and define the function  $r(t)$  by

$$r(t) := \Gamma(d) \int_{p \in \mathbb{R}_{\geq 0}^d} \exp(-\chi \sum_{i=1}^d \lambda_i p_i) \delta(t - \sum_{i=1}^d p_i) \prod_{i=1}^d dp_i. \quad (\text{A.3.8})$$

such that  $r(1) = \mathcal{L}\{f_X\}(\chi)$ . Applying a second Laplace transform to the function  $r(t)$  sending  $t \mapsto \tau$ , produces

$$\mathcal{L}\{r\}(\tau) = \Gamma(d) \prod_{i=1}^d \int_{p_i \in \mathbb{R}_{\geq 0}} \exp(-(\chi\lambda_i + \tau)p_i) dp_i. \quad (\text{A.3.9})$$

As this integration is now separable with respect to  $p_i$ -coordinates, one obtains (assuming  $\chi\lambda_i + \tau > 0$  for the sake of convergence),

$$\mathcal{L}\{r\}(\tau) = \Gamma(d) \prod_{i=1}^d \frac{1}{\chi\lambda_i + \tau}. \quad (\text{A.3.10})$$

In order to apply the inverse Laplace transform, [Eq. \(A.3.6\)](#), one must identify the poles in  $\mathcal{L}\{r\}(\tau)$  as complex function of  $\tau \in \mathbb{C}$ . It is at this stage that one can invoke the simplifying assumption that  $X$  has non-degenerate spectra meaning  $\lambda_i \neq \lambda_j$  for all  $i \neq j$ . Under this assumption, the function  $\mathcal{L}\{r\}(\tau)$  has only *simple* poles whenever  $\tau = -\chi\lambda_i$  for some  $i \in [d]$ . Therefore by the residue theorem (and closing the contour of integration to the left at  $-\infty$ ), the inverse Laplace transform produces

$$r(t) = \Gamma(d) \sum_{i=1}^d \text{Res} \left( \exp(\tau t) \prod_{j=1}^d \frac{1}{\chi\lambda_j + \tau}, x \mapsto -\chi\lambda_i \right), \quad (\text{A.3.11})$$

$$= \Gamma(d) \sum_{i=1}^d \frac{\exp(-\chi\lambda_i t)}{\chi^{d-1} \prod_{j \neq i} (\lambda_j - \lambda_i)}. \quad (\text{A.3.12})$$

The Laplace transform of  $f_X(x)$  is recovered when the normalization parameter  $t$ , is set back to  $t = 1$ , i.e.  $\mathcal{L}\{f_X\}(\chi) = r(1)$ . Therefore,  $f_X(x)$  can be recovered by another application of [Eq. \(A.3.6\)](#). Now  $\mathcal{L}\{f_X\}(\chi)$  has a single pole of order  $d - 1$  at  $\chi = 0$ , so we obtain by an analogous calculation (which contour of integration closed to the right) that

$$f_X(x) = \Gamma(d) \sum_{i=1}^d \frac{\text{H}(x - \lambda_i)}{\prod_{j \neq i} (\lambda_j - \lambda_i)} \text{Res} \left( \frac{\exp(\chi(x - \lambda_i))}{\chi^{d-1}}, \chi \mapsto 0 \right), \quad (\text{A.3.13})$$

where the factor of  $\text{H}(x - \lambda_i)$  (with  $\text{H}$  the Heaviside step function) arises because if  $x - \lambda_i$  were negative, the resulting contour of integration excludes the only pole at  $\chi = 0$ , and thus the integral vanishes. Moreover, the value of the residue itself is simply

$$\text{Res} \left( \frac{\exp(\chi(x - \lambda_i))}{\chi^{d-1}}, \chi \mapsto 0 \right) = \frac{1}{(d-2)!} \lim_{\chi \rightarrow 0} \partial_{\chi}^{(d-2)} \exp(\chi(x - \lambda_i)), \quad (\text{A.3.14})$$

$$= \frac{1}{(d-2)!} (x - \lambda_i)^{d-2}. \quad (\text{A.3.15})$$

In conclusion, the above analysis proves the following theorem which is equivalent to [\[Zha+22, Prop. 4\]](#) (or under the substitution  $2\text{H}(t) = \text{sign}(t) + 1$  to [\[VZ13, Eq. \(17\)\]](#)).

**Theorem A.3.1.** *Let  $X$  be a Hermitian operator on a  $d$ -dimensional Hilbert space  $\mathcal{H}$  with non-degenerate spectra  $\lambda = \{\lambda_1, \dots, \lambda_d\}$ . Then the probability density,  $f_X(x)$ , of the push-forward of the uniform measure on  $\mathbb{P}\mathcal{H}$  through the expectation-value map  $\psi \mapsto \text{Tr}(P_\psi X)$  is given by a piece-wise polynomial function of  $x$ :*

$$f_X(x) = (d-1) \sum_{i=1}^d \frac{(x - \lambda_i)^{d-2} \mathbf{H}(x - \lambda_i)}{\prod_{j \neq i} (\lambda_j - \lambda_i)}. \quad (\text{A.3.16})$$

Unfortunately, [Theorem A.3.1](#) is limited to the special case of non-degenerate spectra, which excludes examples where  $X$  is a projector operator or when  $X$  is a local operator having the form  $L \otimes I$ .

Nevertheless, generalizing [Theorem A.3.1](#) to cover the case of degenerate spectra, while tedious, is fairly straight-forward. Specifically, the only modification that needs to be made in the above derivation is to recognize that degeneracies in the spectra of  $X$  give rise to non-simple poles appearing in [Eq. \(A.3.10\)](#). The proof of the following generalization of [Theorem A.3.1](#) can be found in [[VZ13](#), Sec. 3 & App. B].

**Theorem A.3.2.** *Let  $X$  be a Hermitian operator on a  $d$ -dimensional complex Hilbert space  $\mathcal{H}$  with  $\ell$  distinct eigenvalues  $\{\lambda_1, \dots, \lambda_\ell\}$  where the eigenvalue  $\lambda_j$  has multiplicity (or degeneracy)  $d_j$  where  $1 \leq d_j \leq d$  and  $\sum_{j=1}^{\ell} d_j = d$ . Then the probability density  $f_X(x)$ , as described in [Theorem A.3.1](#), has the form*

$$\begin{aligned} f_X(x) = & \Gamma(d) \sum_{k=1}^{\ell} \sum_{M_k=0}^{d_k-1} \frac{(\lambda_k - x)^{d+M_k-d_k-1} \text{sign}(\lambda_k - x) (-1)^{M_k}}{2(d+M_k-d_k-1)! (d_k-1-M_k)!} \times \\ & \times \sum_{\substack{\{m_j\}_{j=1}^{\ell} \\ \sum_{j \neq k} m_j = M_k}} \prod_{j \neq k} \binom{d_j + m_j - 1}{m_j} \frac{1}{(\lambda_k - \lambda_j)^{d_j + m_j}}. \end{aligned} \quad (\text{A.3.17})$$

**Remark A.3.3.** One might wonder how to generalize the above derivation to handle the case of multiple observables. For instance, if  $A$  and  $B$  are two Hermitian observables, then the joint probability density  $f_{A,B}(a, b)$  induced by pushing forward the uniform measure  $\mathbb{P}\mathcal{H}$  through the map  $\psi \mapsto (\text{Tr}(P_\psi A), \text{Tr}(P_\psi B))$  is

$$f_{A,B}(a, b) = \int_{\psi \in \mathbb{P}(\mathbb{C}^d)} \delta(a - \text{Tr}(P_\psi A)) \delta(b - \text{Tr}(P_\psi B)) d\mu(\psi). \quad (\text{A.3.18})$$

Interestingly one can relate the joint probability density  $f_{A,B}(a, b)$  for the pair of observables  $A$  and  $B$  to the probability density  $f_{\alpha A + \beta B}$  associated to the linear combination  $\alpha A + \beta B$ . To accomplish this, one utilizes the linearity of  $\text{Tr}(P_\psi \cdot)$  and the Laplace transform of the Dirac distributions  $\delta(a - \text{Tr}(P_\psi A))$  and  $\delta(b - \text{Tr}(P_\psi B))$  to obtain the joint Laplace transform of  $f_{A,B}$ :

$$\mathcal{L}\{f_{A,B}\}(\alpha, \beta) = \int_{\psi \in \mathbb{P}(\mathbb{C}^d)} \exp(-\text{Tr}(P_\psi(\alpha A + \beta B))) d\mu(\psi). \quad (\text{A.3.19})$$

However, in order to calculate the above expression using the techniques introduced in this section, it appears necessary to derive an expression for the eigenvalues of  $\alpha A + \beta B$  as a



function of the pair  $(\alpha, \beta) \in \mathbb{R}^2$ . If the operators  $A$  and  $B$  commute,  $[A, B] = 0$ , then this problem becomes straightforward as the eigenvalues of linear combinations of  $A$  and  $B$  are simply linear combinations of the eigenvalues of  $A$  and  $B$  respectively. If however,  $A$  and  $B$  do not commute,  $[A, B] \neq 0$ , then this problem becomes far too difficult, except in the case of low-dimensional Hilbert spaces where relatively simple formulas for roots of characteristic polynomials actually exist.

**Example A.3.4.** The joint probability density for a pair of qubit expectation values,  $A$  and  $B$ , in the sense of [Remark A.3.3](#), can be found in [[Zha+22](#), Prop. 5]. Assuming  $A$  and  $B$  are both traceless and linearly independent qubit observables, the expression in [[Zha+22](#), Prop. 5] simplifies to

$$f_{A,B}(a, b) = \frac{\mathbb{H}(1 - \omega_{A,B}(a, b))}{2\pi \sqrt{\det(T_{A,B})(1 - \omega_{A,B}^2(a, b))}}, \quad (\text{A.3.20})$$

where  $T_{A,B}$  is the invertible matrix

$$T_{A,B} = \begin{pmatrix} \langle A, A \rangle_{\text{HS}} & \langle A, B \rangle_{\text{HS}} \\ \langle B, A \rangle_{\text{HS}} & \langle B, B \rangle_{\text{HS}} \end{pmatrix} \quad (\text{A.3.21})$$

and where  $\omega_{A,B}(a, b) = \sqrt{(a, b)T_{A,B}^{-1}(a, b)^T}$ . This formula was used previously to derive the joint density found in [Eq. \(1.4.7\)](#).

## A.4 A quantum de Finetti theorem

The following theorem, originally [[Chr+07](#), Thm. II.2], demonstrates that the marginals of states belonging to the irreducible representation space  $\mathcal{H}_{\mu+\nu}$  of  $\text{U}(d)$  (when viewed as bipartite states with respect to the inclusion of  $\mathcal{H}_{\mu+\nu}$  in  $\mathcal{H}_{\mu} \otimes \mathcal{H}_{\nu}$ ) are well-approximated by convex combinations of coherent states.

**Lemma A.4.1.** *Let  $\lambda$  be a highest weight in the irreducible representation space  $\mathcal{H}_{\lambda}$  of the compact Lie group  $G$  containing the highest-weight vector  $v_{\lambda} \in \mathcal{H}_{\lambda}$ . For each  $g \in G$ , let the twirled highest-weight vector be  $v_{\lambda}^g = \Phi_{\lambda}(g)v_{\lambda} \in \mathcal{H}_{\lambda}$ , and define the rank-one quasi-projection operator  $E_{\lambda}^g$  by*

$$E_{\lambda}^g := \dim \mathcal{H}_{\lambda} |v_{\lambda}^g\rangle\langle v_{\lambda}^g|. \quad (\text{A.4.1})$$

*Then  $E_{\lambda}^g$  is the density of a positive-operator valued measure with respect to the  $G$ -invariant Haar measure  $dg$ , i.e.*

$$\int_{g \in \text{U}(d)} dg E_{\lambda}^g = \mathbb{1}_{\mathcal{H}_{\lambda}}. \quad (\text{A.4.2})$$

**Lemma A.4.2.** *Let everything be defined as in [Lemma A.4.1](#) and let  $G = \text{U}(d)$  where the weights  $\lambda$  can be identified with integer partitions with at most  $d$  parts or equivalently Young diagrams with at most  $d$  rows.*

*If  $\mu$  and  $\nu$  are integer partitions, then the multiplicity of the representation  $\mathcal{H}_{\mu+\nu}$  inside the inner product representation  $\mathcal{H}_{\mu} \otimes \mathcal{H}_{\nu}$  is exactly one and highest weight vector  $v_{\mu+\nu}$  in  $\mathcal{H}_{\mu+\nu}$  can be identified with the product  $v_{\mu} \otimes v_{\nu}$ . In other words,*

$$E_{\mu+\nu}^g = \frac{\dim(\mathcal{H}_{\mu+\nu})}{\dim(\mathcal{H}_{\mu}) \dim(\mathcal{H}_{\nu})} E_{\mu}^g \otimes E_{\nu}^g. \quad (\text{A.4.3})$$



**Theorem A.4.3.** *Let everything be as in Lemma A.4.2 and let  $\rho_{\mu,\nu} \in \mathcal{S}(\mathcal{H}_{\mu+\nu}) \subseteq \mathcal{S}(\mathcal{H}_\mu \otimes \mathcal{H}_\nu)$  be a bipartite quantum state and let  $\rho_\mu \in \mathcal{S}(\mathcal{H}_\mu)$  be the partial trace over  $\mathcal{H}_\nu$  of  $\rho_{\mu,\nu}$ . Then there exists a probability measure,  $m$ , over  $U(g)$  and a convex combination of coherent states  $C_m^\mu \in \mathcal{S}(\mathcal{H}_\mu)$ , such that*

$$\|\rho_\mu - C_m^\mu\| \leq 4\epsilon, \quad (\text{A.4.4})$$

where  $\|\cdot\|$  is the trace norm ( $\|X\| = \text{Tr}(\sqrt{X^*X})$ ), and

$$\epsilon = 1 - \frac{\dim(\mathcal{H}_\nu)}{\dim(\mathcal{H}_{\mu+\nu})}. \quad (\text{A.4.5})$$

*Proof.* Here we present the same proof technique used in [Chr+07, Thm. II.2]. For ease of notation, let  $P_\lambda^g = |v_\lambda^g\rangle\langle v_\lambda^g|$  be the rank-one projection operator satisfying  $E_\lambda^g = \dim(\mathcal{H}_\lambda)P_\lambda^g$ . Now define  $w_g$  by

$$w_g = \text{Tr}[(\mathbb{1}_\mu \otimes E_\nu^g)\rho_{\mu,\nu}] \quad (\text{A.4.6})$$

and the residual state  $\rho_\mu^g \in \mathcal{S}(\mathcal{H}_\mu)$  by

$$\rho_\mu^g = w_g^{-1} \text{Tr}_\nu[(\mathbb{1}_\mu \otimes E_\nu^g)\rho_{\mu,\nu}] \quad (\text{A.4.7})$$

such that

$$\rho_\mu = \int_{U(d)} dg w_g \rho_\mu^g. \quad (\text{A.4.8})$$

Also, let the probability measure  $m$  on  $U(d)$  be defined such that

$$C_m^\mu = \int_{U(d)} dm(g) P_\mu^g, \quad (\text{A.4.9})$$

$$= \int_{U(d)} \text{Tr}[E_{\mu+\nu}^g \rho_{\mu,\nu}] P_\mu^g dg, \quad (\text{A.4.10})$$

$$= \frac{\dim(\mathcal{H}_{\mu+\nu})}{\dim(\mathcal{H}_\nu)} \int_{U(d)} \text{Tr}[(P_\mu^g \otimes E_\nu^g)\rho_{\mu,\nu}] P_\mu^g dg, \quad (\text{A.4.11})$$

$$= \frac{\dim(\mathcal{H}_{\mu+\nu})}{\dim(\mathcal{H}_\nu)} \int_{U(d)} w_g \text{Tr}[P_\mu^g \rho_\mu^g] P_\mu^g dg, \quad (\text{A.4.12})$$

$$= \frac{\dim(\mathcal{H}_{\mu+\nu})}{\dim(\mathcal{H}_\nu)} \int_{U(d)} w_g P_\mu^g \rho_\mu^g P_\mu^g dg, \quad (\text{A.4.13})$$

where the last equality follows because  $P_\mu^g$  is rank-one. Now using the above integral expressions for  $\rho_\mu$  and  $C_m^\mu$  along with the definition of  $\epsilon$ , the difference  $\rho_\mu - C_m^\mu$  can be expressed as

$$\rho_\mu - C_m^\mu = \rho_\mu - \frac{\dim(\mathcal{H}_\nu)}{\dim(\mathcal{H}_{\mu+\nu})} C_m^\mu - \epsilon C_m^\mu = \int_{U(d)} dg w_g (\rho_\mu^g - P_\mu^g \rho_\mu^g P_\mu^g) - \epsilon C_m^\mu. \quad (\text{A.4.14})$$

At this stage, note that the operator  $\rho_\mu^g - P_\mu^g \rho_\mu^g P_\mu^g$  need not be positive semidefinite. Nevertheless, by letting the maximal projector orthogonal to  $P_\mu^g$  be denoted by  $P_\mu^{-g} = \mathbb{1}_\mu - P_\mu^g$ , one obtains

$$\rho_\mu^g - P_\mu^g \rho_\mu^g P_\mu^g = \rho_\mu^g P_\mu^{-g} + P_\mu^{-g} \rho_\mu^g - P_\mu^{-g} \rho_\mu^g P_\mu^{-g}. \quad (\text{A.4.15})$$

Then, by the definition of  $w_g \rho_\mu^g$  given above,

$$\int_{\mathbf{U}(d)} dg w_g P_\mu^g \rho_\mu^g = \int_{\mathbf{U}(d)} dg \text{Tr}_\nu [(P_\mu^g \otimes E_\nu^g) \rho_{\mu,\nu}] = \frac{\dim(\mathcal{H}_\nu)}{\dim(\mathcal{H}_{\mu+\nu})} \rho_\mu, \quad (\text{A.4.16})$$

which, in turn, implies the contribution of the  $P_\mu^{-g} \rho_\mu^g$  term is of order  $\epsilon$ :

$$\int_{\mathbf{U}(d)} dg w_g P_\mu^{-g} \rho_\mu^g = \epsilon \rho_\mu. \quad (\text{A.4.17})$$

In a similar manner, one can also prove the contribution from the  $\rho_\mu^g P_\mu^{-g}$  is the same, as

$$\int_{\mathbf{U}(d)} dg w_g \rho_\mu^g P_\mu^{-g} = \epsilon \rho_\mu. \quad (\text{A.4.18})$$

Finally, using the convexity of the trace norm and cyclicity of the trace yields

$$\left\| \int_{\mathbf{U}(d)} dg w_g P_\mu^{-g} \rho_\mu^g P_\mu^{-g} \right\| \leq \int_{\mathbf{U}(d)} dg w_g \text{Tr}(P_\mu^{-g} \rho_\mu^g P_\mu^{-g}) = \epsilon \text{Tr}(\rho_\mu). \quad (\text{A.4.19})$$

Altogether,

$$\|\rho_\mu - C_m^\mu\| \leq 3\epsilon \text{Tr}(\rho_\mu) + \epsilon \text{Tr}(C_m^\mu) = 4\epsilon. \quad (\text{A.4.20})$$

□

## A.5 The Laplace principle

An alternative, yet ultimately equivalent approach, to the subject of large deviation theory is to consider it as a generalization of Laplace's method for approximating integrals of exponentials of bounded continuous functions  $g : X \rightarrow \mathbb{R}$  over an interval  $X = [a, b] \subset \mathbb{R}$ . Recall that Laplace's method in this setting yields the expression

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \int_a^b dx \exp(ng(x)) = \max_{a \leq x \leq b} g(x). \quad (\text{A.5.1})$$

The following condition generalizes this idea to the setting where (i) the domain of integration,  $X$ , is more general, and (ii) the measure over  $X$  used in the integration depends on the value of  $n$ .

**Definition A.5.1** (Laplace Principle). A sequence of probability measures  $(\mu_n : \Sigma(X) \rightarrow [0, 1])_{n \in \mathbb{N}}$  on a standard Borel space  $(X, \Sigma(X))$  satisfies the **Laplace principle** with rate function  $I : X \rightarrow [0, \infty]$  if for all bounded continuous functions  $g \in \mathcal{C}_b(X)$ ,

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \int_{x \in X} \mu_n(dx) \exp(ng(x)) = \sup_{x \in X} (g(x) - I(x)). \quad (\text{A.5.2})$$

That the large deviation principle implies the Laplace principle (with the same rate function) is known as Varadhan's integral lemma [Var66] (see also [DE11, Thm. 1.2.1] or [DZ10, Sec. 4.3]). Also note that a converse to Varadhan's lemma also holds [DE11, Thm. 1.2.3]. In order to see the connection between these two approaches more carefully, we will first need to better understand how the rate function is directly related to the asymptotics of the sequence of probability measures  $(\mu_n)_{n \in \mathbb{N}}$ .

**Definition A.5.2.** Let  $(X, \Sigma(X))$  be a standard Borel space, let  $\mu : \Sigma(X) \rightarrow [0, 1]$  be a probability measure, and let  $g \in X \rightarrow \mathbb{R}$  be a measurable function. Define the **(generalized) cumulant generating function** for  $\mu$ , denoted by  $\Lambda$ , as

$$\Lambda(g) := \log \int_X \mu(dx) \exp(g(x)). \quad (\text{A.5.3})$$

If  $(\mu_n : \Sigma(X) \rightarrow [0, 1])_{n \in \mathbb{N}}$  is a sequence of probability measure on  $X$ , then define the **(generalized) regularized<sup>3</sup> cumulant generating function** as

$$\Lambda_n(g) := \frac{1}{n} \log \int_X \mu_n(dx) \exp(ng(x)), \quad (\text{A.5.4})$$

and let  $\Lambda_\infty(g)$  be the result of the following limit, if it exists:

$$\Lambda_\infty(g) = \lim_{n \rightarrow \infty} \Lambda_n(g). \quad (\text{A.5.5})$$

**Remark A.5.3.** The function  $\Lambda$  defined above is referred as the *generalized* cumulant generating function simply because if  $g_1 : X \rightarrow \mathbb{R}$  and  $g_2 : X \rightarrow \mathbb{R}$  are measurable functions corresponding to independent random variables,  $X_1$  and  $X_2$ , meaning for all regions  $\Delta_1, \Delta_2 \in \Sigma(\mathbb{R})$ ,

$$((g_1, g_2)_* \mu)(\Delta_1, \Delta_2) = ((g_1)_* \mu)(\Delta_1) ((g_2)_* \mu)(\Delta_2), \quad (\text{A.5.6})$$

or perhaps expressed more commonly as,

$$p(X_1 \in \Delta_1, X_2 \in \Delta_2) = p(X_1 \in \Delta_1) p(X_2 \in \Delta_2), \quad (\text{A.5.7})$$

then  $\Lambda$  is *cumulative* in the sense that

$$\Lambda(g_1 + g_2) = \Lambda(g_1) + \Lambda(g_2). \quad (\text{A.5.8})$$

Moreover, if the special case where  $X$  happens to be a vector space and the measurable function  $g : X \rightarrow \mathbb{R}$  is linear, i.e.

$$g(x) = \langle \lambda, x \rangle \quad (\text{A.5.9})$$

for some  $\lambda$ , then the cumulant function  $\Lambda(\lambda) := \Lambda(\langle \lambda, \cdot \rangle)$  becomes the usual cumulant generating function for the probability measure  $\mu$  (see [DZ10, Pg. 26]).

The next result, known as Bryc's theorem, expresses conditions under which it becomes possible to prove that a sequence probability measures satisfies the large deviation principle, and moreover what the rate function must be equal to [DZ10, Thm. 4.4.2].

---

<sup>3</sup>Instead of the adjective *regularized*, some authors prefer the adjective *scaled* or *upper-limiting* [DZ10; Tou11].

**Theorem A.5.4** (Bryc’s theorem). *Let  $(X, \Sigma(X))$  be a standard Borel space. Let  $(\mu_n : \Sigma(X) \rightarrow [0, 1])_{n \in \mathbb{N}}$  be a sequence of probability measures that is exponentially tight (Definition 2.1.23). If for all bounded, continuous functions,  $g \in X \rightarrow \mathbb{R} \in \mathcal{C}_b(X)$ , the limit  $\Lambda_\infty(g)$  of the regularized cumulant generating function (see Eq. (A.5.5)) for  $(\mu_n)_{n \in \mathbb{N}}$  exists, then  $(\mu_n)_{n \in \mathbb{N}}$  satisfies the large deviation principle with rate function*

$$I(x) := \sup_{g \in \mathcal{C}_b(X)} (g(x) - \Lambda_\infty(g)). \quad (\text{A.5.10})$$

Furthermore, for every  $g \in \mathcal{C}_b(X)$ ,

$$\Lambda_\infty(g) = \sup_{x \in X} (g(x) - I(x)). \quad (\text{A.5.11})$$

*Proof.* For a proof, see [DZ10, Thm. 4.4.2], [DE11, Thm. 1.3.8], or the Bryc’s original paper [Bry90].  $\square$

In light of Theorem A.5.4, specifically Eq. (A.5.11), it makes sense to formalize the following transformation of rate functions.

**Definition A.5.5.** Let  $X$  be a set, let  $I : X \rightarrow [0, \infty]$  be a function and let  $g \in \mathcal{C}_b(X)$  be a bounded continuous real-valued function on  $X$ . We define the **(generalized) Fenchel-Legendre transform** of  $I$  by

$$I^*(g) := \sup_{x \in X} (g(x) - I(x)). \quad (\text{A.5.12})$$

**Remark A.5.6.** If the set  $X$  is a vector space and the function  $g$  is taken to be a linear function on  $X$ , i.e.  $g(x) = \langle \lambda, x \rangle$ , then the map sending  $\lambda$  to

$$I^*(\langle \lambda, \cdot \rangle) = \sup_{x \in X} (\langle \lambda, x \rangle - I(x)), \quad (\text{A.5.13})$$

is better known as the *Fenchel-Legendre transform* of  $I$  [DZ10]. When the function  $I$  is additionally convex, then  $I^*$  is known as the *Legendre transform* of  $I$ .

**Remark A.5.7.** Given definitions for the regularized cumulant generating function (Definition A.5.2) and the generalized Fenchel-Legendre transform (Definition A.5.5), we see that a sequence of probability measures  $(\mu_n : \Sigma(X) \rightarrow [0, 1])_{n \in \mathbb{N}}$  satisfies the Laplace principle with rate function  $I : X \rightarrow [0, \infty]$  if and only if the limit of the regularized cumulant generating function for  $(\mu_n)_{n \in \mathbb{N}}$  exists and equals the Fenchel-Legendre transform of  $I$  for all bounded continuous real-valued functions on  $X$ :

$$\forall g \in \mathcal{C}_b(X) : \Lambda_\infty(g) = \lim_{n \rightarrow \infty} \Lambda_n(g) = I^*(g). \quad (\text{A.5.14})$$

## A.6 Asymptotic Born rule

The following result can be interpreted as saying that if you measure an unknown quantum source  $n$  times with binary projective measurement  $\{P, I - P\}$  and obtain the outcome associated to  $P$  each time, then in the limit of large  $n$ , the probability of obtaining outcome  $Q$  after applying the projective measurement  $\{Q, I - Q\}$  is given by the Born rule for the density operator  $\rho_P$  corresponding to the normalization of  $P$ .

**Lemma A.6.1.** Let  $P, Q \in \text{End}(\mathbb{C}^d)$  be Hermitian operators where  $P^2 = P$ , let  $\rho_P = P/\text{Tr}(P)$  be a density operator associated to  $P$  and let  $n \in \mathbb{N}$ . Then

$$\text{Tr}(\rho_P Q) = \lim_{n \rightarrow \infty} \frac{\text{Tr}(\sum_{n+1}(P^{\otimes n} \otimes Q))}{\text{Tr}(\sum_n(P^{\otimes n}))}. \quad (\text{A.6.1})$$

*Proof.* A direct calculation yields

$$\text{Tr}(\sum_{n+1}(P^{\otimes n} \otimes Q)) = \frac{(d-1)! (n+d_P-1)!}{(n+d)! (d_P-1)!} [\text{Tr}(Q) + n\text{Tr}(\rho_P Q)]. \quad (\text{A.6.2})$$

As a sanity check, note when  $n = 0$ ,  $\text{Tr}(\sum_1(Q)) = \frac{1}{d}\text{Tr}(Q)$ , and when  $Q = I$ ,

$$\text{Tr}(\sum_{n+1}(P^{\otimes n} \otimes I)) = \frac{(d-1)! (n+d_P-1)!}{(n+d)! (d_P-1)!} (d+n) = \frac{\binom{n+d_P-1}{n}}{\binom{n+d-1}{n}}, \quad (\text{A.6.3})$$

as expected. In any case, since  $\text{Tr}(\sum_n(P^{\otimes n})) = \text{Tr}(\sum_{n+1}(P^{\otimes n} \otimes I))$ , we conclude

$$\frac{\text{Tr}(\sum_{n+1}(P^{\otimes n} \otimes Q))}{\text{Tr}(\sum_n(P^{\otimes n}))} = \frac{\text{Tr}(Q) + n\text{Tr}(\rho_P Q)}{d+n}. \quad (\text{A.6.4})$$

In the limit as  $n \rightarrow \infty$ , only the  $\text{Tr}(\rho_P Q)$  remains and thus the claim holds.  $\square$

**Remark A.6.2.** Note that if the operator  $Q$  in [Lemma A.6.1](#) is a non-zero projective effect (meaning  $0 < Q \leq I$  and  $Q^2 = Q$ ), then the expression in [Eq. \(A.6.4\)](#) is minimized when  $Q$  is both rank one and orthogonal to the density operator  $\rho_P$  in which case the minimum probability assignable to the effect  $Q$  conditioned on  $n$  repeated observations of the effect  $P$  is given by  $(d+n)^{-1}$ , i.e.

$$\frac{\text{Tr}(\sum_{n+1}(P^{\otimes n} \otimes Q))}{\text{Tr}(\sum_n(P^{\otimes n}))} \geq \frac{1}{d+n}. \quad (\text{A.6.5})$$

This bound is consistent with the claim made by Blume-Kohout that the minimum probability that can be assigned to an element of a  $d$ -outcome measurement after  $n$  trials is  $(d+n)^{-1}$  [[Blu10](#), Eq. (2)].