

Resource theories of quantum coherence: foundations and applications



Benjamin Yadin
Wolfson College
University of Oxford

A thesis submitted for the degree of
Doctor of Philosophy

Hilary 2017

Abstract

One of the fundamental features that separates quantum physics from classical physics is the idea of quantum superposition, also known as coherence. This thesis concentrates on understanding quantum coherence in the mathematical framework of resource theories, viewing it both as a resource to be harnessed and as a way to quantitatively characterise quantum states in contrast to classical states.

We first examine the type of coherence resource theory which has emerged recently to cope with general settings where the physical nature of the medium encoding information is not crucial, such as computation. We identify the set of quantum processes in which coherence is neither created nor used, and use these to provide a physically motivated resource theory pictured in terms of interferometry. Using the same concepts, we then find connections between coherence and discord, a type of quantum correlation. In particular, we show how coherence can be used to generate discord, and explore basis-dependent discord as an intermediate quantity.

The second part of the thesis applies the resource theory framework to quantify quantum macroscopicity, taken here to mean the extent to which coherence exists in a system on a macroscopic scale. We find the appropriate type of resource theory for this purpose, giving criteria for good measures of macroscopic coherence. We use these criteria to evaluate some previously proposed measures and highlight the role of the quantum Fisher information. Next, we build up measures based on the concept of macroscopic distinguishability and use them to show that macroscopic quantum states are fragile to noise induced by interaction with an environment. Finally, we apply measures based on the Fisher information to a range of experiments involving mechanical degrees of freedom, in order to compare their macroscopicity.

Acknowledgements

I firstly owe many thanks to my supervisor, Vlatko Vedral, who has been a great source of guidance, encouragement and inspiration. I am also extremely grateful to Davide Girolami for both introducing me to the subject of quantum coherence and helping direct my research. All the members of the group in Oxford have helped create a friendly and stimulating environment that made the past years thoroughly enjoyable. In particular, the more senior students – Andrew Garner, Nana Liu, and the Felixes (or Felices) Pollock, Binder, and Tennie – never hesitated to pass on their knowledge and insight. I would further like to thank the remaining past and current members: Tristan Farrow, Oscar Dahlsten, Chiara Marletto, Christian Schilling, Fabio Anzà, Pieter Bogaert, Anu Unnikrishnan, Tian Zhang, Reevu Maity, and Cormac Browne. Very many thanks to Thomas Veness and Thomas Elliott for their valued friendship since undergraduate times.

Many thanks to Mile Gu for being so welcoming during visits to his group, and for his valuable advice, in collaboration alongside Jiajun Ma. I also thank Florian Fröwis for hosting a productive visit to Geneva and for many resulting discussions. For proof-reading, I am particularly indebted to Felix Tennie, Reevu, and Davide. I gratefully acknowledge funding from the EPSRC and Wolfson College in connection with the scholarship from Simon Harrison.

Finally, I can't thank my family enough, in particular my brother and parents, for their continued support and generosity, and especially Vera for managing to put up with me every day.

Contents

1	Overview	1
2	Introduction	6
2.1	Quantum states and observables	6
2.1.1	Basics	6
2.1.2	Qubits	9
2.1.3	Continuous variables	10
2.2	Quantum operations	12
2.2.1	Closed systems – unitary evolution	12
2.2.2	Projective measurements	13
2.2.3	Gates and circuits	14
2.2.4	General quantum operations	14
2.2.5	Master equations	16
2.3	Information theory – classical and quantum	17
2.3.1	Entropy	17
2.3.2	Distinguishability	18
2.3.3	Fisher Information	20
2.4	Resource theories	21
2.4.1	Basic structure	21
2.4.2	State transformations	23
2.4.3	Measures of resources	24
2.5	Quantum macroscopicity	25
2.5.1	Model-independent tests	26
2.5.2	Collapse models	27
2.5.3	Quantum properties	27

3	Free operations for speakable coherence	29
3.1	Incoherent operations – first attempt	30
3.2	Incoherent operations which do not use coherence	32
3.3	Physical meaning of SI operations	35
3.3.1	Operations in an interferometer	35
3.3.2	Dilation	37
3.4	Monotones	40
3.5	Comparison with other incoherent operations	43
3.5.1	Translationally invariant operations	43
3.5.2	Genuinely and fully incoherent operations	45
3.5.3	Physical incoherent operations	48
3.5.4	Dephasing-covariant incoherent operations	49
3.5.5	Coherent transport and cooling maps	51
3.6	Conclusions	52
4	Relating coherence and discord	55
4.1	Basics of quantum discord	56
4.2	Creating discord from coherence	59
4.2.1	Measures of multipartite coherence and discord	59
4.2.2	General bounds	62
4.2.3	Application to DQC1 algorithm	66
4.3	Incoherent operations and discord	68
4.3.1	Basis-dependent discord (BDD)	68
4.3.2	Vanishing BDD	70
4.3.3	Behaviour under local operations	73
4.3.4	Recovery from local dephasing	78
4.3.5	Implications for basis-independent discord	81
4.4	Conclusions	86
5	Quantifying macroscopic coherence	88
5.1	Defining the problem	88
5.2	Digression: asymmetry and reference frames	90

5.3	Monotones	92
5.3.1	Distance-based measures	92
5.3.2	Variance and QFI	93
5.3.3	Generalised QFI	95
5.4	QFI as an asymptotic preparation cost	95
5.4.1	Continuous spectrum	96
5.4.2	Harmonic spectrum	97
5.4.3	Arbitrary discrete spectrum	98
5.4.4	Formation of mixed states	99
5.5	From coherence to macroscopicity	100
5.5.1	Measure by Fröwis and Dür	100
5.5.2	Measure by Lee and Jeong	103
5.5.3	Measure by Nimmrichter and Hornberger	107
5.6	Conclusions	109
6	Macroscopic distinguishability	111
6.1	Measurement model	112
6.2	Construction	114
6.3	Properties	117
6.3.1	Coherence criteria	117
6.3.2	Limiting cases	119
6.4	Alternative measure using discord	122
6.5	Fragility of micro-macro entangled states	127
6.6	Conclusions	129
7	Macroscopic coherence and entanglement of mechanical systems	131
7.1	Two elements of macroscopicity	132
7.1.1	General considerations	132
7.1.2	Extensive size	133
7.1.3	Entangled size	134
7.2	Estimates for physical systems	137
7.2.1	Leggett’s crystal	137
7.2.2	Molecular diffraction	139

7.2.2.1	Far-field estimates	140
7.2.2.2	Near-field estimates	141
7.2.3	Mechanical oscillators	141
7.2.4	Proposed quantum gravity test	142
7.2.5	Hypothetical Schrödinger’s cat	143
7.3	Conclusions	143
8	Conclusions and outlook	146
A	Generalised QFI	150
A.1	Background	150
A.2	Deterministic monotone	152
A.3	Ensemble monotone	152
B	N_{LJ} decoherence model	153
C	Details of monotonicity of dist_Δ	155
D	Degradation of entanglement of formation	156
E	Quadrature QFI and variance in a thermal state	157
	Bibliography	158

Chapter 1

Overview

During the development of quantum theory in the early twentieth century, it became evident that certain distinguishing features set the new theory far apart from its classical predecessor. The “quantisation” that provides the theory with its name refers to the fact that certain quantities, such as energy, have to be discrete in situations where they would classically be continuous. This was the foundation of Einstein’s explanation of the photoelectric effect [8] in terms of particles of light (photons) and the Bohr-Sommerfeld model of the atom [9,10]. On the other hand, as hypothesised by de Broglie and first experimentally shown for electrons by Davisson and Germer [11], objects that previously would have been envisaged as discrete particles also have a wave-like character. As had been understood much earlier, the observation of interference in classical optics requires coherence. In general (there are spatial and temporal versions), coherence means that different parts of a wave have a fixed phase relationship [12]. The same applies to interference of quantum particles, although now the phase is an intrinsic property of the wavefunction, rather than a classical field. If quantum coherence exists between two points in the wavefunction of a particle, then it can be understood as being in a superposition of both locations simultaneously.

More generally, quantum coherence in an arbitrary system can be taken to mean the existence of a superposition of at least two different states. Note that quantum coherence is understood to mean something subtly different (though definitely related) in quantum optics, referring to correlations between fields at different points that cannot exist with classical light [13,14]. While the concept of coherence, and various measures to quantify it, have been ubiquitous for decades, it is only recently that a full rigorous mathematical framework for axiomatising such measures in great generality

has been sought [15].

The framework at its core regards quantum coherence as a resource – something which can be exploited to perform a task. This perspective has a long and successful history throughout the development of quantum information theory and quantum technologies, in which genuinely quantum features are utilised as resources to provide an advantage (such as greater speed, efficiency or precision) over classical resources. The plethora of developing and potential quantum technologies could be said to originate with Feynman’s visionary proposal that the high complexity of quantum processes is best simulated by quantum systems rather than classical computers [16,17]. The work of Deutsch [18] and others similarly led to the subject of quantum computing, whose aim is to identify quantum algorithms for a range of computational problems [19]. While a universal quantum computer has not yet been realised, a number of systems including superconducting devices, trapped ions, and photons are promising candidates [20]. Beyond computing, there are possibilities for using quantum systems in cryptography [21–23] and other communication tasks [24]. Quantum metrology is the science of using quantum states as probes of physical quantities or parameters in order to attain a level of statistical precision beyond classical limits [25–27]. Most recently, there has been theoretical and experimental evidence to suggest that coherence plays a crucial role in accounting for the efficiency of some biological processes, including energy transport in photosynthesis and magneto-reception in birds [28–32].

Of all quantum resources, arguably none has been explored more than entanglement [33]. It is known to be a resource in (pure state) quantum computation [34], dense coding [35] and state teleportation [36], to name just a few cases. The mathematical structures for rigorously defining measures of entanglement were developed with its role as a resource in mind [37,38]. This has led to the pervasive view that fundamental quantum properties can be understood and quantified by thinking about how they might be useful. Therefore the motivation for a resource theory is not only practical: we expect that this approach can tell us something deeper about the structure of quantum states. For example, considerations about reversibility have uncovered parallels between entanglement theory and the second law of thermodynamics [39].

In Chapter 3, we consider how to build a resource theory of coherence of a type first suggested recently by Baumgratz et al. [40]. This has been termed a theory

of “speakable” coherence by Marvian and Spekkens [41], referring to its envisaged applicability in situations typically considered as “information-theoretic”, where the exact nature of the physical system encoding information is not important. The strength of this proposal is its attempt at great generality; its weakness is that its physical relevance is not immediately clear. We propose a modification of the original resource theory which we argue properly takes into account the resources needed to perform operations in which coherence is not created. We give a physical basis for this theory by relating it to interferometry, and make a detailed comparison with similar recent proposals.

In Chapter 4, we connect this theory of coherence to a kind of quantum correlation known as discord [42–44]. We show how the coherence of individual subsystems can be traded for discord between subsystems. Furthermore, we give some new perspectives on the status of discord as a resource, including a critical analysis of whether either coherence or discord is the crucial resource for a noisy model of computation [45]. In addition, while discord has no known resource theory, we offer some insights for potential progress in this direction by studying its behaviour under processes where coherence is valuable.

In the latter part of the thesis, we change perspective to view coherence as a way of distinguishing quantum states from classical states, in particular at large scales. Thanks to extensive experimental testing of quantum theory to high precision, the superposition principle has been shown to hold at least for microscopic systems. In principle, there is nothing in the theory that prevents this from being true up to the macroscopic scale. However, that is at odds with our experience of the classical macroscopic world. The apparent absurdity of macroscopic superpositions was famously highlighted by Schrödinger’s thought experiment [46] of a cat inside a box, whose fate is determined by a nucleus decaying probabilistically via a coherent quantum process. Before being disturbed by an outside entity such as an observer, if quantum theory holds fully, the state of the contents of the box should be a superposition of two “branches” which are macroscopically distinct (a dead cat versus a live cat).

One way of explaining why such states are never seen is to realise that the box is never truly isolated: there is always an interaction with the environment. Such interactions typically result in decoherence, whereby quantum effects are diluted. The

need for increasingly high precision in the control of ever larger and more complex systems, in order to avoid decoherence, is responsible for the great experimental challenge of quantum technologies [47, 48]. Similar arguments attribute the difficulty of observing macroscopic superpositions to limited precision in measurements [49–51]. However, arguments of this type do not explain why we only see one branch of the superposition. A number of more extreme proposals posit new physical laws, including fundamental modifications to quantum mechanics and possible effects due to quantum gravity [52–54]. These predict negligible changes to the dynamics of microscopic systems but a fast collapse of macroscopic superpositions.

The matter can only be settled by experimental tests – in recent decades, quantum superpositions have been pushed to ever larger scales [55–57]. While none of these could yet be described as macroscopic, the term mesoscopic might be appropriate. Early breakthroughs were made in the demonstration of wave-like behaviour of C_{60} molecules [58], and of the superposition of electrical currents differing by the order of μA in superconducting devices [59–61]. Since then, progress has been made in these systems and in others, including photonic systems [62, 63] and micro-mechanical oscillators [64].

With such a variety of systems and types of states, it is desirable to have a way of quantifying and comparing how well each experiment tests the superposition principle at the macroscopic scale. In the literature, a large number of such measures have been proposed [54, 55, 65–70] – typically these are referred to measures of macroscopic quantumness or quantum “macroscopicity”. This proliferation of measures, for which it is often hard to see a relation between any two, shows that deciding on a consistent definition of a “macroscopic quantum state” is no trivial task.

Chapter 5 proposes a general formal structure for deciding what quantities are appropriate for measuring macroscopicity. This is another resource theory of coherence – of the “unspeakable” type, in that the physical basis for the information storage is crucial. We give criteria for good measures of macroscopic coherence and test some previous proposals against these criteria, highlighting the importance of one measure based on the quantum Fisher information.

In Chapter 6, we explore the concept of macroscopic distinguishability phrased in terms of the measurement precision needed to discriminate between branches of a superposition. We construct coherence measures from this principle and show that

they satisfy the resource theory criteria of Chapter 5. As a consequence, we find new results about the susceptibility of entanglement to decoherence when a microscopic system is coupled to a macroscopic one.

Finally, in Chapter 7, we analyse a range of experiments involving superpositions of states which are distinct with respect to position or momentum variables, in particular, diffracting molecules and mechanical oscillators. We find two complementary measures for such systems – one which is extensive the system size, and another which measures the number of atoms which are genuinely entangled – and compare their values in the experiments.

Publications

All the results and presentation here are the author’s own, although often developed in collaboration with others. It is stated explicitly where a result or idea is entirely due to another person but included for completeness. The work in Chapters 3 and 4 is a rearrangement of the papers [3, 4], with some additional results. Chapter 5 is based on [2], and Chapter 6 is a collection of the parts of [6] which were derived by the author. Chapter 7 is based on a manuscript in preparation.

Chapter 2

Introduction

The purpose of this chapter is to introduce the background material that is relevant for understanding the context and main concepts of the research presented throughout the thesis. We start here with the basic machinery of quantum mechanics and quantum information – this is mainly to establish terminology and notation. Next, we introduce the general structure of quantum resource theories and the types of quantities they usually deal with. Finally, we give some background to the subject of quantifying quantum macroscopicity, with an overview of the types of approaches taken and measures proposed in previous literature. Some more specialised and technical background also appears in each of the later chapters. In particular, the fundamentals of coherence resource theories are discussed at the start of Chapter 3 as a way of building up to the main point of the chapter.

Some general remarks about notation and conventions: we write operators without “hats”, for example Q instead of $\hat{Q}$ – the only exception is the bosonic mode number operator $\hat{n}$ which we distinguish from the integer variable n . After this introduction, we also use natural units in which $\hbar = k_B = 1$, apart from in Chapter 7 where the SI unit values of quantities become important. We use the term “iff” as an abbreviation for “if and only if”.

2.1 Quantum states and observables

2.1.1 Basics

In this thesis, we adopt the Dirac notation [71] for quantum mechanics that has become standard among much of the physics community. The reader is referred to

standard texts such as [72,73]. In a closed quantum system, full knowledge about the state of the system is described by a state vector, or **ket** $|\psi\rangle$ which is an element of a Hilbert space $\mathcal{H}$, often referred to as a **state space**. For a given physical state, there is a global phase freedom under which $|\psi\rangle$ and $e^{i\theta}|\psi\rangle$ are equivalent for any $\theta \in \mathbb{R}$. To every $|\psi\rangle$ there corresponds a unique dual vector, or **bra**, $\langle\psi|$, such that $\langle\psi|\phi\rangle$ is the Hilbert space inner product between $|\psi\rangle$ and $|\phi\rangle$. While the **amplitude** $\langle\psi|\phi\rangle$ is a complex number, $|\langle\psi|\phi\rangle|^2 \geq 0$ and is interpreted as the probability of obtaining a “yes” outcome if we perform a measurement that asks “is the system in state $|\psi\rangle$ after being prepared in $|\phi\rangle$?” – this is known as the **Born rule**. Such an interpretation is only possible if we impose a normalisation constraint $\langle\psi|\psi\rangle = 1$ for all state vectors. Unless otherwise stated explicitly, we will assume that every state vector is normalised.

Two orthogonal states (satisfying $\langle\psi|\phi\rangle = 0$) can be regarded as mutually exclusive and thus fully distinguishable. Given a **basis** $\{|i\rangle\}$ consisting of mutually orthogonal states (i.e. $\langle i|j\rangle = \delta_{i,j}$, using the Kronecker delta) which span the entire state space $\mathcal{H}$, every state can be expressed in the form

$$|\psi\rangle = \sum_i c_i |i\rangle, \quad c_i \in \mathbb{C}, \quad (2.1)$$

where the normalisation constraint is equivalent to $\sum_i |c_i|^2 = 1$. (Although we have used discrete, possibly infinite, sums here, there are situations in which an integral is appropriate – see Section 2.1.3.) The corresponding expansion for the dual vector is $\langle\psi| = \sum_i c_i^* \langle i|$.

While $|c_i|^2$ is the probability of measuring $|\psi\rangle$ to be in state $|i\rangle$, we are emphatically not to think of the **superposition** (2.1) as the description of a system of whose state we are merely uncertain. Indeed, the use of the complex amplitudes c_i in place of probabilities is often argued to be the fundamental difference between quantum and classical physics [74]. That adding amplitudes is quite distinct from adding probabilities can be seen from the following example. Let there be two basis states $|0\rangle, |1\rangle$ and define

$$|\pm\rangle := \frac{1}{\sqrt{2}}(|0\rangle \pm |1\rangle). \quad (2.2)$$

Both $|+\rangle$ and $|-\rangle$ have an equal probability of being found in each of $|0\rangle$ and $|1\rangle$. However, their superposition is $(|+\rangle + |-\rangle)/\sqrt{2} = |0\rangle$ – this answer would not be obtained by using the rules of classical probability. This phenomenon is at the heart of quantum interference and wave-particle duality.

In order to accommodate classical probabilities into the formalism, we represent a general quantum state by a **density operator** ρ , which is a linear operator on $\mathcal{H}$ that satisfies $\rho^\dagger = \rho$ (i.e., is hermitian), $\rho \geq 0$, and the normalisation condition $\text{Tr } \rho = 1$. A pure state $|\psi\rangle$ corresponds to the outer product $\rho = |\psi\rangle\langle\psi|$, while a classical probabilistic **mixture** of a set of pure states $\{|\psi_\alpha\rangle\}$ with probabilities $\{p_\alpha\}$ is

$$\rho = \sum_{\alpha} p_{\alpha} |\psi_{\alpha}\rangle\langle\psi_{\alpha}|. \quad (2.3)$$

A sum of this form – as contrasted with a superposition such as (2.1) – will be referred to here as a **mixture**, and may in some places be called a convex sum. For a given ρ , such a decomposition is typically not unique, and in fact can be done in infinitely many different ways [72]. In a measurement which asks if ρ is in a state $|\phi\rangle$, the probability of “yes” is $\langle\phi|\rho|\phi\rangle$ – which is consistent with both the Born rule and the classical probabilities of (2.3).

Let us now describe how to treat composite systems. For two subsystems A and B with respective Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$, the state space of the combined system is the tensor product $\mathcal{H}_A \otimes \mathcal{H}_B$. Similarly, two completely independent states $|\psi\rangle_A, |\phi\rangle_B$ are described in the combined system by the tensor product $|\psi\rangle_A |\phi\rangle_B := |\psi\rangle_A \otimes |\phi\rangle_B$. An arbitrary state of AB will typically be written as a sum of terms of this type. Likewise, for density operators, two independent states ρ_A, σ_B are represented by $\rho_A \otimes \sigma_B$.

The density operator formalism is particularly useful for describing smaller parts of a larger system. For a state ρ_{AB} of a bipartite system AB , the state of A on its own is described by the **reduced density operator**

$$\rho_A := \text{Tr}_B \rho_{AB} = \sum_i \langle i |_B \rho_{AB} | i \rangle_B, \quad (2.4)$$

where the sum can be over any basis of $\mathcal{H}_B$.

Finally, any physically measurable quantity in quantum mechanics is associated with an **observable**, which is represented by a hermitian operator on $\mathcal{H}$. Writing an observable Q in a spectral decomposition, $Q = \sum_i q_i |i\rangle\langle i|$, where each eigenvector $|i\rangle$ is a state for which measuring Q yields the value q_i . From the Born rule, it can be seen that the **expectation value** of Q (i.e., its statistical average from a measurement in

the $\{|i\rangle\}$ basis) in a state ρ reads

$$\langle Q \rangle_\rho := \sum_i q_i \langle i | \rho | i \rangle = \text{Tr}(Q\rho). \quad (2.5)$$

In a composite system (say, AB), an operator acting non-trivially on a single part A is of the form $Q_A \otimes I_B$ – we will often write this as Q_A for brevity.

A **projector** Π is a particular kind of observable which corresponds to a “yes/no” measurement – mathematically, it is characterised by $\Pi^2 = \Pi$, and so has eigenvalues 0 and 1. When an observable Q is degenerate, each distinct eigenvalue q is associated with a projector $\Pi_q = \sum_{i:q_i=q} |i\rangle\langle i|$. Thus we can also write $Q = \sum_q q \Pi_q$. These projectors are orthogonal, meaning that $\Pi_q \Pi_{q'} = \Pi_q \delta_{q,q'}$. They are also complete, such that $\sum_q \Pi_q = I$.

In general, two projectors Π, Π' might not commute, so their **commutator** $[\Pi, \Pi'] := \Pi\Pi' - \Pi'\Pi$ is non-zero. Physically, this means that the two measurements cannot simultaneously have definite outcomes in every state. Similarly, two non-commuting observables cannot always have simultaneously well-defined values. One way to express this quantitatively is via the Heisenberg-Robertson uncertainty principle [75]

$$V(\rho, A)V(\rho, B) \geq \frac{1}{4} |\langle [A, B] \rangle_\rho|^2, \quad (2.6)$$

where $V(\rho, A) := \langle A^2 \rangle_\rho - \langle A \rangle_\rho^2$ is the variance (a measure of uncertainty) of A in the state ρ .

2.1.2 Qubits

The simplest quantum system is the qubit, which has a Hilbert space of dimension two. It is standard to refer to the basis $\{|0\rangle, |1\rangle\}$ as the computational basis. Along with the identity operator, the **Pauli operators** provide a basis for all observables; written as matrices in the computational basis, these are

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \sigma^x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma^y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma^z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (2.7)$$

Every qubit state can be written in the form $\rho = \frac{1}{2}(I + \sum_{k=x,y,z} r_k \sigma^k)$, where $\mathbf{r} = (r_x, r_y, r_z)$ is a real vector of modulus $|\mathbf{r}| \leq 1$. Considering $\mathbf{r}$ as a direction in Euclidean space, the state space can be pictured as a ball of unit radius centred at the origin – the **Bloch sphere**. The set of pure states corresponds to the surface of the sphere, while the **maximally mixed state** $\rho = I/2$ lies at the origin.

The computational basis is also the eigenbasis of σ^z , and the states lying on the z -axis consist of all mixtures $p|0\rangle\langle 0| + (1-p)|1\rangle\langle 1|$ where $p \in [0, 1]$. By contrast, the eigenstates of σ^x , $|\pm\rangle$, and of σ^y , $(|0\rangle \pm i|1\rangle)/\sqrt{2}$, are superpositions of the computational basis states.

2.1.3 Continuous variables

Please see a reference such as [76] for a full exposition of this subject. First we outline the mathematical tools suitable for describing the (non-relativistic) motion of massive particles. A single particle of mass m has position and momentum variables $\mathbf{x}, \mathbf{p}$ satisfying the commutation relation $[x_i, p_j] = i\hbar\delta_{i,j}$ ($i = 1, 2, 3$ for three-dimensional space) – we refer to these variables as **quadratures**. In free space, its Hamiltonian is $H = \frac{\mathbf{p}^2}{2m}$. Let us concentrate for now on one spatial dimension; when there is a harmonic trapping potential of frequency ω , the Hamiltonian becomes

$$H = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 x^2. \quad (2.8)$$

By introducing the **annihilation and creation operators** $a, a^\dagger$, this may be rewritten as

$$H = \hbar\omega \left(a^\dagger a + \frac{1}{2} \right), \quad \text{where} \quad a := \frac{m\omega x + ip}{\sqrt{2m\hbar\omega}}. \quad (2.9)$$

These satisfy the **bosonic** commutation relation $[a, a^\dagger] = 1$; with multiple modes a_i in play, one would have $[a_i, a_j^\dagger] = \delta_{i,j}$. The **number operator** $\hat{n} := a^\dagger a$ has eigenstates $|n\rangle$, $n = 0, 1, 2, \dots$ which each contain a definite number of vibrational quanta. In particular, the **vacuum state** $|0\rangle$ has no quanta but still has **zero-point motion**, referring to the fact that $V(|0\rangle\langle 0|, x)$ and $V(|0\rangle\langle 0|, p)$ are non-zero, so the particle cannot be thought of as standing still. The creation operator can be used to construct all number states from the vacuum state via

$$|n\rangle = \frac{a^{\dagger n}}{\sqrt{n!}} |0\rangle. \quad (2.10)$$

As photons are bosonic particles, the same machinery is used in quantum optics – photonic modes are mathematically described as harmonic oscillators, only with electromagnetic field components in place of position and momentum. For a single mode, the Hamiltonian has the same form $H = \hbar\omega(a^\dagger a + \frac{1}{2})$, where ω is the mode frequency. Number states are described in this context as having a definite number of

photons. The quadratures are formally the same as above, but are now related linearly to components of the electric and magnetic fields; we express them in dimensionless form as¹

$$x = \frac{1}{\sqrt{2}}(a + a^\dagger), \quad p = \frac{1}{i\sqrt{2}}(a - a^\dagger). \quad (2.11)$$

Apart from the number states, there are some other families of states that are frequently encountered in both mechanical and optical systems. A **coherent state**² with the parameter $\alpha \in \mathbb{C}$ is defined by the eigenvalue equation $a|\alpha\rangle = \alpha|\alpha\rangle$. By introducing the **displacement operator** $D(\alpha)$, this can be related to the vacuum state:

$$|\alpha\rangle = D(\alpha)|0\rangle, \quad \text{where} \quad D(\alpha) := e^{\alpha a^\dagger - \alpha^* a}. \quad (2.12)$$

The word “displacement” stems from the action on the quadratures

$$D(\alpha)^\dagger x D(\alpha) = x + \sqrt{2}\text{Re}(\alpha), \quad D(\alpha)^\dagger p D(\alpha) = p + \sqrt{2}\text{Im}(\alpha). \quad (2.13)$$

Note that the average photon number is given by $\langle n \rangle = |\alpha|^2$. Large-amplitude coherent states (and mixtures thereof) are typically considered as good models of classical light. Part of the reason for this is that the (quantum) fluctuations in number ($O(\sqrt{\langle n \rangle})$) and in the quadratures ($O(1)$) are negligible compared with the expected photon number $\langle n \rangle$. Coherent laser light may be modelled by a pure coherent state for which $\arg \alpha$ is the fixed phase of the laser.

Similarly, the **squeezing operator** $S(\xi)$ lets us construct **quadrature-squeezed coherent states**

$$|\alpha; \xi\rangle := D(\alpha)S(\xi)|0\rangle, \quad \text{where} \quad S(\xi) := e^{\frac{1}{2}(\xi^* a^2 - \xi a^{\dagger 2})}. \quad (2.14)$$

In the case $\xi \in \mathbb{R}$, we have $S(\xi)^\dagger x S(\xi) = e^{-\xi}x$, $S(\xi)^\dagger p S(\xi) = e^\xi p$, corresponding to a shrinking scale in one quadrature and an expansion in the conjugate direction (changing the phase of ξ simply rotates the axes of squeezing). Squeezed light may be generated by passing coherent laser light through a non-linear medium involving a process called degenerate parametric down-conversion.

¹Note that different conventions exist in the literature, with varying factors of 2 or $\sqrt{2}$.

²It is unfortunate that this long-established terminology conflicts with our later usage of “coherent state” as meaning “any state that is not incoherent”. In quantum optics, the point is to emphasise that $|\alpha\rangle$ is a superposition of number states. Matters are made worse by our assessment in Chapter 5 that the states $|\alpha\rangle$ can be regarded in a sense as the *least coherent* pure states.

When a (generic) system with Hamiltonian H is in contact with a heat bath at temperature T , and has come to equilibrium, it is described by a **thermal state**

$$\rho_{th} := \frac{e^{-\beta H}}{\text{Tr } e^{-\beta H}}, \quad \beta := \frac{1}{k_B T}. \quad (2.15)$$

For a bosonic mode, this implies that

$$\rho_{th} = \sum_{n=0}^{\infty} \frac{e^{-\beta \hbar \omega (n+1/2)}}{Z} |n\rangle \langle n|, \quad (2.16)$$

where $Z = \sum_{n=0}^{\infty} e^{-\beta \hbar \omega (n+1/2)}$.

Finally, we introduce a representation of continuous-variable systems in phase space. Classically, the state of a system is specified by a point in phase space (which uses positions and momenta as its coordinates) – or in general by a probability distribution over such points. In quantum mechanics, since x and p are not simultaneously measurable, it is impossible for a system to exist at an exact point in phase space, and there is some inherent uncertainty in its location. One way (amongst a number of alternatives) to picture a quantum state in phase space is with the **Wigner function**, also known as a **quasi-probability distribution**. This can be defined via an integral over position-space matrix elements as

$$\mathcal{W}_\rho(q, s) := \frac{1}{2\pi\hbar} \int_{-\infty}^{\infty} dr \, e^{irs/\hbar} \langle q + r/2 | \rho | q - r/2 \rangle. \quad (2.17)$$

Its marginals recover the probability distribution in any quadrature, for example $\int_{-\infty}^{\infty} ds \, \mathcal{W}_\rho(q, s) = \langle q | \rho | q \rangle$ (where $|q\rangle$ is a position eigenstate). However, unlike a probability distribution, $\mathcal{W}_\rho$ can be negative. For example, the Wigner function of a coherent state $|\alpha\rangle$, $\alpha = (q_0 + is_0)/\sqrt{2}$ is a Gaussian $\mathcal{W}(q, s) = \frac{2}{\pi} e^{-(q-q_0)^2 - (s-s_0)^2}$, which is positive everywhere. On the other hand, for a number state $|n\rangle$, we have $\mathcal{W}(q, s) = \frac{2}{\pi} (-1)^n L_n(2q^2 + 2s^2) e^{-q^2 - s^2}$ (where L_n is the n th Laguerre polynomial), which is much more oscillatory and goes below zero. See Fig. 2.1 for plots of these two cases.

2.2 Quantum operations

2.2.1 Closed systems – unitary evolution

Having discussed states at single instances in time, let us now focus on how systems evolve. It is a postulate of quantum mechanics that the state of a closed system

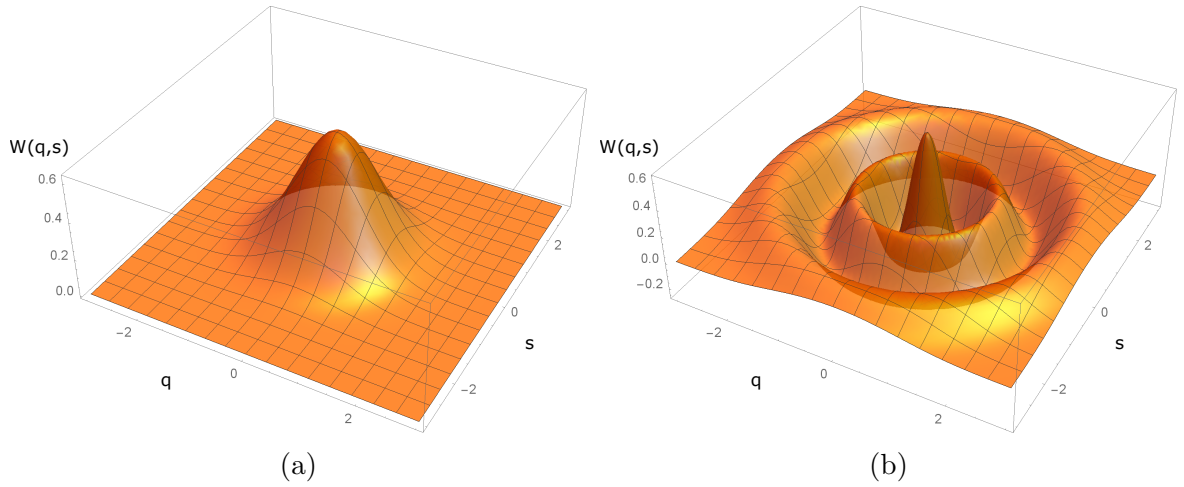


Figure 2.1: Wigner function plots for (a) the vacuum state $|0\rangle$ and (b) a number state $|3\rangle$.

changes from one time to another by a **unitary** transformation, such that $|\psi\rangle \rightarrow |\psi'\rangle = U|\psi\rangle$, where $UU^\dagger = U^\dagger U = I$. At the level of density operators, this is written as $\rho \rightarrow \rho' = U\rho U^\dagger$. The **time-dependent Schrödinger equation** defines how this evolution is generated by the Hamiltonian operator H , the observable corresponding to energy:

$$i\hbar \frac{d|\psi\rangle}{dt} = H|\psi\rangle \quad \text{or} \quad i\hbar \frac{d\rho}{dt} = [H, \rho]. \quad (2.18)$$

Assuming that H is time-independent, the unitary operator giving evolution over a time t is $U_t = e^{-itH/\hbar}$; we say that H **generates** the evolution U_t .

2.2.2 Projective measurements

We have already described how to calculate the probabilities of outcomes of simple “yes/no” measurements, but have not stated how such a measurement affects the state. This is the content of a separate postulate of quantum mechanics. Let us state this for a measurement defined by a complete set of mutually commuting projectors $\{\Pi_i\}$: the outcome i occurs with probability $p_i = \text{Tr}(\Pi_i \rho)$ and results in the transformation

$$\rho \rightarrow \rho'_i = \frac{\Pi_i \rho \Pi_i}{p_i}, \quad (2.19)$$

where the denominator ensures that ρ_i is normalised. An experimenter who knows that a measurement has been performed, but is ignorant of the outcome, will instead

describe the final state as the mixture

$$\rho \rightarrow \rho' = \sum_i p_i \rho'_i = \sum_i \Pi_i \rho \Pi_i. \quad (2.20)$$

2.2.3 Gates and circuits

In the subject of quantum computation and algorithms (with which we make brief contact in Chapter 4), discrete unitary transformations are referred to as **gates**. Sequences of gates in combination with measurements are known as **circuits**, and may be depicted in diagrams using wires to represent subsystems and boxes to represent operations – the reader is directed to a standard text such as [72] for the conventions of such diagrams. Common single-qubit gates include the Pauli operators σ^α , as well as the **Hadamard** gate,

$$H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \quad (2.21)$$

written in the computational basis. Note that σ^x can also be referred to as a NOT gate, since its action on the computational basis is the same as a classical NOT gate, i.e. $\sigma^x |0\rangle = |1\rangle$, $\sigma^x |1\rangle = |0\rangle$. Interesting computations also require gates that interact subsystems. One such entangling gate is the two-qubit controlled-NOT, or **C-NOT**, which can be written as

$$|0\rangle\langle 0| \otimes I + |1\rangle\langle 1| \otimes \sigma^x. \quad (2.22)$$

This performs a σ^x gate on the second (or “target”) qubit based on the state of the first (or “control”) qubit.

2.2.4 General quantum operations

The main ingredients of unitary evolution and projective measurements as defined above can be combined into a formalism describing a large class of processes in quantum systems. These are so general and ubiquitous in descriptions of open-system dynamics that they are typically referred to as **quantum operations**. We first provide the physical picture, and next give the mathematical details. Suppose we have a system S and an ancilla α (which may play the role of an environment when the operation describes some kind of noise). The initial state is uncorrelated, i.e. $\rho_S \otimes \sigma_\alpha$, and undergoes closed-system dynamics described by some unitary U . Upon discarding

the ancilla, the resulting transformation on the system is

$$\rho_S \rightarrow \mathcal{E}(\rho_S) := \text{Tr}_\alpha[U(\rho_S \otimes \sigma_\alpha)U^\dagger]. \quad (2.23)$$

The operation $\mathcal{E}$ is called a **channel**, and is easily seen to be a linear transformation in the sense that $\mathcal{E}(\sum_i p_i \rho_i) = \sum_i p_i \mathcal{E}(\rho_i)$ for any set of states ρ_i and probabilities p_i . It is also **trace-preserving**, such that $\text{Tr} \mathcal{E}(\rho) = \text{Tr} \rho$, and **positive**, so that $\mathcal{E}(X) \geq 0$ for any operator $X \geq 0$. These conditions are physically important, as they ensure that $\mathcal{E}$ is consistent with classical probabilistic mixing and that it preserves the set of physical states. Moreover, $\mathcal{E}$ is **completely positive (CP)**: for any additional system R and a positive operator X_{RS} , $\mathcal{E}_S(X_{RS}) \geq 0$.³ A stronger condition than positivity, this ensures that $\mathcal{E}$ keeps a state physical even when acting only on part of it.

A channel can be very usefully represented by a set of **Kraus operators** [77] $\{K_\mu\}$:

$$\mathcal{E}(\rho) = \sum_\mu K_\mu \rho K_\mu^\dagger, \quad \text{where} \quad \sum_\mu K_\mu^\dagger K_\mu = I. \quad (2.24)$$

Importantly, a linear operation is trace-preserving and CP iff it has a Kraus operator representation (2.24), and iff there exists a state σ_α and unitary U providing a representation (2.23) – referred to as a **dilation** for the channel [78, 79]. Note that it is always sufficient to take σ_α to be a pure state, say $|0\rangle$. Given a basis $\{|\phi_\mu\rangle\}$ for $\mathcal{H}_\alpha$, a set of Kraus operators can be constructed by

$$K_\mu = \langle \phi_\mu |_\alpha U | 0 \rangle_\alpha. \quad (2.25)$$

For a single channel, the Kraus operator representation is not generally unique: the dilation is fixed up to a change of basis on α , and correspondingly each basis choice gives a different set of Kraus operators.

A channel is a deterministic operation; by allowing for CP maps that are **trace-non-increasing**, i.e. $\text{Tr} \mathcal{E}(\rho) \leq \text{Tr} \rho$, we can account for operations that are **non-deterministic**. Then $\text{Tr} \mathcal{E}(\rho)$ is the probability that $\mathcal{E}$ is successfully applied to ρ . In terms of Kraus operators, we just alter (2.24) to the condition $\sum_\mu K_\mu^\dagger K_\mu \leq I$. In the dilation picture, this corresponds to making a projective measurement on α instead of tracing it out.

³We use $\mathcal{E}_S$ as shorthand for $I_R \otimes \mathcal{E}_S$, where I_R is the identity operation on R .

For any CP map $\mathcal{E}(\rho) = \sum_{\mu} K_{\mu} \rho K_{\mu}^{\dagger}$, we can define the **adjoint** map by conjugating the Kraus operators: $\mathcal{E}^{\dagger}(\rho) = \sum_{\mu} K_{\mu}^{\dagger} \rho K_{\mu}$.

As another variant, let us suppose that a measurement on α is made by a complete set of projectors $\{\Pi_{\mu}\}$. Selecting each outcome μ leads to a generally non-deterministic operation $\mathcal{E}_{\mu}$. However, the sum $\sum_{\mu} \mathcal{E}_{\mu}$ is a channel, and results from losing all information about which outcome was obtained. We refer to the set $\{\mathcal{E}_{\mu}\}$ as a **quantum instrument** [80]. In general, many different instruments can be compatible with the same channel. In the case where each $\mathcal{E}_{\mu}$ has a single Kraus operator K_{μ} , each selected outcome occurs with probability $p_{\mu} = \text{Tr}(K_{\mu}^{\dagger} K_{\mu} \rho)$ and has the effect

$$\rho \rightarrow \rho'_{\mu} = \frac{K_{\mu} \rho K_{\mu}^{\dagger}}{p_{\mu}}, \quad (2.26)$$

which is a generalisation of the projective measurement (2.19). The set of observables $E_{\mu} := K_{\mu}^{\dagger} K_{\mu}$ is said to define a kind of generalised measurement known as a **positive operator-valued measure (POVM)**.

Note that all the above statements may be extended to account for operations that have different input and output state spaces. We will not make use of this generalisation here – but note that the simple cases of tracing out a subsystem and adding an uncorrelated subsystem can also be treated as quantum channels. Another way of having a different output space is to perform a quantum instrument with a **classical memory** [81]. For this, the memory X is modelled as an additional system which encodes a label of the outcome μ into a set of orthogonal basis states $\{|\mu\rangle\}$. Thus the instrument $\{\mathcal{E}_{\mu}\}$ on a system S is associated with a channel

$$\mathcal{F}_S(\rho) = \sum_{\mu} |\mu\rangle\langle\mu|_X \otimes \mathcal{E}_{\mu_S}(\rho). \quad (2.27)$$

2.2.5 Master equations

Instead of using discretely applied operations, one can describe many noise processes in continuous time by using a **master equation**. This is a modification of the time-dependent Schrödinger equation in the following way:

$$\frac{d\rho}{dt} = -\frac{i}{\hbar} [H, \rho] + \sum_{\mu} 2L_{\mu} \rho L_{\mu}^{\dagger} - \{L_{\mu}^{\dagger} L_{\mu}, \rho\}, \quad (2.28)$$

where $\{A, B\} := AB + BA$ is the anticommutator, and the L_{μ} are known as Lindblad operators [82]. Such a description is appropriate for describing a system interacting

with an environment, under the Born and Markov approximations, which state that the system and environment only become weakly correlated, and that the environment has no memory effects [83]. Common cases include environment-induced decoherence, for instance from dephasing processes, and dissipation, for example from spontaneous emission.

2.3 Information theory – classical and quantum

2.3.1 Entropy

One of the most fundamental quantities in classical information theory is the entropy introduced by Shannon [84]. This is defined for a discrete probability distribution $\{p_i\}$ by⁴

$$H(\{p_i\}) := - \sum_i p_i \log p_i. \quad (2.29)$$

This attains a minimum of zero for distributions having unit probability of a single outcome. For a distribution with support on a set of size d , the maximum of $\log d$ is achieved by the uniform distribution $p_i = 1/d \forall i$. Hence, one can think of the Shannon entropy as measuring the uncertainty that an observer has about a random variable when they know its distribution but not its value. In thermodynamics, it takes on physical relevance in the form of the Gibbs entropy (which is essentially the same) [85]; Shannon gave it an operational interpretation as the optimal compression rate for a signal composed of many identically distributed and independent instances of a random variable. In Chapter 6 we will make use of the **differential entropy** $H(p(x))$, defined for a continuous random variable by replacing p_i with the density $p(x)$, and the sum with an integral. One must take some care with interpreting this, however, as it can be negative [86]. Let us also mention an alternative notation: if a random variable X has values x_i with probabilities p_i , we can write its Shannon entropy as $H(X)$.

The analogous quantity for quantum states is the **von Neumann entropy**

$$S(\rho) := - \text{Tr}(\rho \log \rho) = H(\{\lambda_i\}), \quad (2.30)$$

where the second equality shows that this is the same as the Shannon entropy of the eigenvalues of ρ . This is non-negative, zero iff ρ is pure, and, on a Hilbert space of

⁴Unless stated explicitly otherwise, we will always take the logarithm with base 2.

dimension d , maximised at $\log d$ for the maximally mixed state $\rho = I/d$. It is then plausible that $S(\rho)$ measures the degree of “mixedness” of ρ . Along similar lines, the quantity $\text{Tr}(\rho^2)$ is often called the **purity** of ρ , being maximised at unity for pure states and minimised at $1/d$ for maximally mixed states. Both $S(\rho)$ and $\text{Tr}(\rho^2)$ are invariant under unitary transformations $\rho \rightarrow U\rho U^\dagger$. Similarly to the classical case, the von Neumann entropy was first given an interpretation as a compression rate in a communication protocol by Schumacher [87].

2.3.2 Distinguishability

Given two probability distributions $\{p_i\}, \{q_i\}$, one can ask how to quantify how distinct they are. We give three notable examples: these are the **trace distance**

$$D_{tr}(\{p_i\}, \{q_i\}) := \frac{1}{2} \sum_i |p_i - q_i|, \quad (2.31)$$

the **fidelity**

$$F(\{p_i\}, \{q_i\}) := \sum_i \sqrt{p_i q_i}, \quad (2.32)$$

and the **relative entropy** (or Kullback-Leibler divergence)

$$H(\{p_i\} || \{q_i\}) := \sum_i p_i (\log p_i - \log q_i). \quad (2.33)$$

All three quantities are non-negative, and for identical distributions (and only for identical distributions), we have $D_{tr}(\{p_i\}, \{p_i\}) = H(\{p_i\} || \{p_i\}) = 0$, $F(\{p_i\}, \{p_i\}) = 1$. Note that the relative entropy is formally defined to be infinite whenever the support of $\{p_i\}$ is not contained in the support of $\{q_i\}$. Each quantity can also be generalised to continuous random variables (and in fact the relative entropy is better behaved than the differential entropy in this situation, being non-negative).

They can all be extended to quantify the distinguishability of two quantum states:

$$D_{tr}(\rho, \sigma) := \frac{1}{2} \|\rho - \sigma\|_{tr}, \quad (2.34)$$

$$F(\rho, \sigma) := \|\sqrt{\rho}\sqrt{\sigma}\|_{tr}, \quad (2.35)$$

$$S(\rho || \sigma) := \text{Tr}(\rho \log \rho - \rho \log \sigma), \quad (2.36)$$

where $\|X\|_{tr} := \text{Tr}(\sqrt{X^\dagger X})$ is the **trace norm** of the operator X . Again note that $S(\rho || \sigma)$ is formally infinite unless $\text{supp } \rho \subseteq \text{supp } \sigma$. Each has a connection to its

classical counterpart, equalling it when ρ and σ are diagonal in the same basis. Furthermore, $D_{tr}(\rho, \sigma)$ equals the maximal classical trace distance between the probability distributions obtained from ρ and σ , optimising over all generalised (POVM) measurements [88]. $F(\rho, \sigma)$ is similarly the minimal classical fidelity over all POVM measurements. The analogous statement for the relative entropy is more complex, relating to the discrimination of $\rho^{\otimes n}$ and $\sigma^{\otimes n}$ in the limit of large n [89].

There is a crucial property of these distinguishability measures: they respect the intuition saying that no noisy operation (i.e., one that gains no new information about the state) should be able to make two states more distinguishable. Stated precisely, for any quantum channel $\mathcal{E}$,

$$D_{tr}(\mathcal{E}(\rho), \mathcal{E}(\sigma)) \leq D_{tr}(\rho, \sigma), \quad F(\mathcal{E}(\rho), \mathcal{E}(\sigma)) \geq F(\rho, \sigma), \quad S(\mathcal{E}(\rho) \| \mathcal{E}(\sigma)) \leq S(\rho, \sigma). \quad (2.37)$$

Using geometrical terminology, these concepts can be generalised to **distance measures**, or **metrics**. There are various properties we might demand of a distance measure D :

- (a) $D(\rho, \sigma) \geq 0$ and $D(\rho, \sigma) = 0$ iff $\rho = \sigma$;
- (b) symmetry, $D(\rho, \sigma) = D(\sigma, \rho)$;
- (c) triangle inequality, $D(\rho, \tau) \leq D(\rho, \sigma) + D(\sigma, \tau)$.

For some cases such as the relative entropy, (b) and (c) are not satisfied – without requiring these, we refer to D as a **pseudo-distance**. We say that D is **contractive** whenever it is non-increasing under an arbitrary channel $\mathcal{E}$ applied to ρ and σ : $D(\mathcal{E}(\rho), \mathcal{E}(\sigma)) \leq D(\rho, \sigma)$. A distance measure satisfying all these conditions can be defined from the fidelity: the **Bures distance**

$$D_B(\rho, \sigma) := \sqrt{2(1 - F(\rho, \sigma))}. \quad (2.38)$$

A class of distance measures (including D_{tr}) may be derived from an operator norm via $D(\rho, \sigma) \propto \|\rho - \sigma\|$ – whether or not this is contractive will depend on the properties of the chosen norm.

2.3.3 Fisher Information

Closely related to the concept of distinguishability is the quantity known as **Fisher information**. Originally defined within classical statistics, this is a central quantity in the subject of parameter estimation. Suppose that there is a family of probability distributions $p_t(x)$ depending on a real parameter t (we assume a continuous distribution here, but most of the statements carry through to the discrete case). In general, there may be multiple parameters, but we only need the single parameter case here. The task is to determine the value of t from a number M of independent samples from the distribution. Denote the estimated value extracted from the measurements by $\tilde{t}$, and assume that this is unbiased in the sense that $\langle \tilde{t} \rangle_t = t$, where the expectation value is taken with respect to $p_t(x)$. The **Cramér-Rao bound** sets a fundamental lower limit on the uncertainty with which t can be estimated:

$$(\Delta t)^2 := \langle (\tilde{t} - t_0)^2 \rangle_{t_0} \geq \frac{1}{M \mathcal{F}_c(p_t(x))|_{t_0}}, \quad (2.39)$$

where the (classical) Fisher information of $p_t(x)$ at $t = t_0$ is defined by

$$\mathcal{F}_c(p_t(x))|_{t_0} := \int_{-\infty}^{\infty} dx \frac{[\partial_t p_t(x)]^2}{p_t(x)} \Big|_{t_0}. \quad (2.40)$$

Importantly, the bound can be asymptotically saturated in the limit of large M . So $\mathcal{F}_c$ is a measure of the achievable accuracy in parameter estimation. One often encounters the case in which the parameter shifts the distribution, so $p_t(x)$ is replaced by $p(x - t)$. Then the translational Fisher information (which is independent of t) is given by

$$\mathcal{F}_c(p(x)) = \int_{-\infty}^{\infty} dx \frac{p'(x)^2}{p(x)}. \quad (2.41)$$

In the subject of quantum metrology [25, 27], a parameter can instead be encoded into a family of states ρ_t . Compared with the classical case, there is an added complication that different measurements will result in different performances. Then quantum Fisher information (QFI) can then be defined by optimising over all POVM measurements; remarkably, this has a closed analytic expression [90]:

$$\mathcal{F}(\rho_t) := \max_{\text{POVM } \{E(x)\}} \mathcal{F}_c(\text{Tr}[E(x)\rho_t]) \quad (2.42)$$

$$= 2 \sum_{i,j} \frac{|\langle \psi_i | \partial_t \rho_t | \psi_j \rangle|^2}{\lambda_i + \lambda_j}, \quad (2.43)$$

where $\rho_t = \sum_i \lambda_i |\psi_i\rangle\langle\psi_i|$ is a spectral decomposition with $\lambda_i, |\psi_i\rangle$ implicitly functions of t . For M independent copies of the state ρ_t , we then have the quantum Cramér-Rao bound

$$(\Delta t)^2 \geq \frac{1}{M \mathcal{F}(\rho_t)|_{t_0}}. \quad (2.44)$$

In the limit of large M , this can also be saturated – the optimal measurement is determined by the **symmetric logarithmic derivative (SLD)** L_t , which may be defined implicitly by $\partial_t \rho_t = \frac{1}{2} \{L_t, \rho_t\}$. The required estimator is then the expectation value of the observable $t_0 + L_{t_0}/\mathcal{F}(\rho_t)_{t_0}$. It is important to note that this depends on t_0 . This means that the QFI measures the performance of a state in *local* parameter estimation, where some information is already known that localises the parameter somewhat, and the task is to determine small fluctuations of t .

Often in metrology, one has N subsystems, each of which can be thought of as a “probe” of some process defined by a channel $\mathcal{E}_t$. If these are all produced independently, then the precision is determined by the QFI of $\mathcal{E}_t(\rho)^{\otimes N}$, which is proportional to N . This then gives the **standard scaling** (sometimes called shot-noise scaling) $\Delta t \sim 1/\sqrt{N}$. However, if the probes can be in a correlated state ρ_N , then the QFI of $(\mathcal{E}_t^{\otimes N})(\rho_N)$ can scale up to N^2 . The resulting **Heisenberg scaling** $\Delta t \sim 1/N$ is often taken as a signature of a quantum advantage in metrology.⁵

Finally, we remark that the QFI is connected to the Bures distance [90]:

$$\mathcal{F}(\rho_t)|_{t_0} = 4\partial_t^2 D_B^2(\rho_{t_0}, \rho_t)|_{t_0}, \quad (2.45)$$

so it can be thought of as the rate at which ρ_t becomes distinguishable from the initial state with increasing t .

2.4 Resource theories

2.4.1 Basic structure

The idea of a resource theory has grown out of retrospectively analysing the general mathematical structures behind well-established areas of science. The formalism takes a practical perspective, aiming to describe what tasks, or transformations of a physical system, can be achieved by an experimenter when certain constraints exist

⁵In quantum optics, it is common to use a single mode and instead define these scaling limits with respect to N being the expected photon number.

on the available states and dynamics. Anything that overcomes these restrictions is considered a resource. Notable examples include thermodynamics, in which systems out of thermal equilibrium can be used to work, and chemistry, where one is interested in the amounts of each reactant required to reach a certain product. Situations as broad as these can indeed be formalised mathematically as resource theories [91].

However, here we are concerned with quantum resource theories, phrased in terms of the basic elements of states and operations. A quantum resource theory is defined once we specify the following:

- the set of **free states** – those which are considered freely available at no cost;
- the set of **free operations** – any quantum operations which respect some given experimental restriction.

In order to arrive at a consistent theory, we must assume that the set of free states is closed under all free operations – otherwise, resource states could be created for free. It also makes sense to include the preparation of all free states in the free operations. Some other assumptions are both typical and often reasonable on the set of free states: it is closed under tensor products, classical mixing, and tracing out subsystems [92]. Note that later we often refer to “the set of free states/operations” as those defined on a single system.

The resource theory approach has been used to study a large variety of quantum phenomena, including purity [93], thermodynamics [94], and steering [95] – not to mention coherence and asymmetry, which we discuss in detail throughout the thesis. These all follow the example of entanglement [39, 96], which is the longest-studied and most successful quantum resource theory (though by no means the simplest).

We therefore use this as an archetypal example of a resource theory; let us define entanglement in a bipartite system AB , although the extension to an arbitrary number of parties is trivial. The first decision to make is the partition between subsystems – this is often naturally given when A and B are spatially distant parties. By definition, a state is said to have no entanglement when it is **separable**, meaning that it can be expressed in the form

$$\rho_{AB} = \sum_i p_i \rho_A^{(i)} \otimes \rho_B^{(i)}. \quad (2.46)$$

For pure states, this reduces to being a product $|\psi\rangle_A|\phi\rangle_B$. The free operations are those that can be performed by two distant labs who individually have full quantum control but can only send each other classical information: these are known as **local operations and classical communication (LOCC)**. This set is very hard to completely describe mathematically [97], but we can at least say that there must be a set of Kraus operators of the product form $K_\mu = L_\mu \otimes M_\mu$. However, the full set of operations with this property are known as **separable operations**, a strictly larger set [38, 98]. The largest possible set, defined solely by the inability to create entanglement from separable states, comprises **non-entangling operations** [39]. LOCC is considered the most practically relevant set, thanks to its clear physical motivation.

2.4.2 State transformations

One of the most fundamental questions one can ask in a resource theory context is whether one state can be transformed into another by using free operations. We write $\rho \rightarrow \sigma$ when there exists a free operation $\mathcal{E}$ such that $\mathcal{E}(\rho) = \sigma$. This is a deterministic transformation; for a stochastic transformation, we can say that $\rho \rightarrow \sigma$ succeeds with some probability p . For example, a major result by Nielsen [99] gives a simple condition for pure state deterministic transformations $|\psi\rangle \rightarrow |\phi\rangle$ under bipartite LOCC.

There may also be a **maximal resource state** from which it is possible to produce any other state (perhaps stochastically). This exists in bipartite entanglement: for two systems of dimension d , a maximally entangled state is of the form $\sum_{i=0}^{d-1} d^{-1/2} |i\rangle|i\rangle$ (using any basis on each party). However, it is also possible for two states to be incomparable, meaning that there is no transformation taking one to the other – either way – with nonzero probability. In such a situation, the two states contain genuinely different resources. This phenomenon occurs in multipartite entanglement [100].

The general mixed state case is typically much harder to deal with, but concrete statements can be made for **asymptotic transformations**. Here, one considers the transformation of a large number N of copies of a state ρ into M copies of σ . Let us suppose that this is deterministic, but might have some error ϵ . Then we write

$$\rho^{\otimes N} \rightarrow \rho'_N \stackrel{\epsilon}{\approx} \sigma^{\otimes M}, \quad \text{i.e.} \quad D_{tr}(\rho'_N, \sigma^{\otimes M}) \leq \epsilon. \quad (2.47)$$

We say that ρ can be asymptotically transformed to σ at a rate R if the above transformation is achievable with $M/N = R$ such that $\epsilon \rightarrow 0$ as $N \rightarrow \infty$.

After designating some chosen (non-free) state Ψ as a “resource unit”, this kind of transformation can be viewed in two ways, as either the preparation of many copies of a given state ρ from resource units, or the distillation of many copies of ρ into units. The **distillation rate** from ρ is defined as the maximal rate R at which ρ is asymptotically transformed to Ψ . Similarly, the **cost** of preparing ρ is the minimal inverse rate $1/R$ at which Ψ can be asymptotically transformed to ρ . For example, in bipartite entanglement theory, the cost E_c and distillation rate E_d utilise the **Bell state**, or “e-bit”, $|\phi^+\rangle = (|0\rangle|0\rangle + |1\rangle|1\rangle)/\sqrt{2}$ as the resource unit. For pure states, both of these quantities reduce to the **entanglement entropy**

$$E_E(|\psi\rangle\langle\psi|_{AB}) := S(\text{Tr}_B |\psi\rangle\langle\psi|_{AB}). \quad (2.48)$$

In general, the distillation rate cannot exceed the cost – this follows because one cannot get more resource units out of a state than used to prepare it. A resource theory for which these quantities are equal is called **asymptotically reversible** – a sufficient condition for this is that the set of free operations is the maximal possible one preserving the set of free states [92]. In contrast, it may happen (as is the case in mixed state entanglement) that some resource states are not distillable. This feature is described as a resource being **bound**.

2.4.3 Measures of resources

Costs and distillation rates are ways to quantify the amount of resource in a state. This idea can be generalised in an axiomatic way to give criteria to be satisfied by any genuine **measure** of a resource. The first axiom simply says that a measure should be a function from states to non-negative real numbers, vanishing iff the state is free.

Next, a measure is required to be non-increasing under free operations – this property is known as **monotonicity**. Hence, we might also refer to good measures as monotones. Fundamentally, this comes from wanting to capture the fact that free operations never add any resource to the system. This point can also be given an operational meaning: we expect a valid measure to give the optimal figure of merit for some task performed within the scope of the resource theory (although there are certainly measures for which such an interpretation is not known). Since the optimal

sequence of operations for the task must be within the set of free operations, incorporating another free operation cannot aid the task further. Costs and distillation rates are clear examples of this. For a full discussion, see [101].

We give the mathematical details and various flavours of monotonicity in context in Chapter 3. We will also see there that other non-essential criteria can be given, such as convexity.

One particularly powerful way of constructing measures is geometrical in nature, quantifying the distance from a given state to the set of free states:

$$\min_{\text{free } \sigma} D(\rho, \sigma). \quad (2.49)$$

By definition, this vanishes exactly on free states. When D is contractive, this measure can be proven to be a monotone under free operations [37].

2.5 Quantum macroscopicity

In order to assess how well a given experiment really probes quantum mechanics at the macroscopic scale, we must first decide exactly what it means for quantum effects to exist at the macroscopic scale. Early thoughts on this matter were expressed by Leggett [55, 65], who made the distinction between genuine macroscopic quantum phenomena and what we might call microscopic quantum effects accumulating to the macroscopic scale. For an example of the latter, consider the stability of solid matter, which certainly manifests itself at the macroscopic scale and cannot be understood with classical mechanics. However, it is explained well by quantum models which extend over microscopic scales with distant atoms being effectively independent. In contrast, an experiment confirming the quantum nature of a macroscopic superposition such as Schrödinger’s cat would require a model treating all parts of the system simultaneously in the quantum formalism.

Very roughly, then, the quantum macroscopicity of a state means the degree to which genuinely quantum effects exist at the macroscopic scale. Of course, this is an extremely vague statement, so there is no surprise that it has been interpreted in a large variety of ways in previous works. In order to identify and motivate a precise approach, we start by categorising past approaches into three main divisions: model-independent tests, collapse models, and quantum properties.

2.5.1 Model-independent tests

Ideally, one would want to measure macroscopicity using as few assumptions as possible about the system and its dynamics. This is arguably desirable if one believes that some new, unknown post-quantum physics is responsible for the macroscopic emergence of classical behaviour.

One approach which has been well studied concerns **Leggett-Garg inequalities (LGIs)** [102, 103], which aim to provide a test of single-system non-classicality along the lines of Bell inequalities [104]. While violation of a Bell inequality rules out local realistic theories, LGIs are derived from two assumptions named **macro-realism** and **non-invasive measurability**. Both of these are properties obeyed by the apparently classical phenomena of the everyday macroscopic world. Macro-realism (MR) means that any macroscopic quantity always has a well-defined value (although an experimenter may have some uncertainty about that value). Non-invasive measurability (NIM) says that it is in principle possible to measure the value of such an observable without disturbing the state. An LGI experiment involves making measurements at different times. For example, suppose the three quantities Q_i ($i = 1, 2, 3$), with possible values ± 1 , are measured at respective times $t_1 < t_2 < t_3$. In any given run of the experiment, only two of these are measured. Assuming MR and NIM, it is straightforward to show the following inequality:

$$\langle Q_1 Q_2 \rangle_{\bar{3}} + \langle Q_1 Q_3 \rangle_{\bar{2}} + \langle Q_2 Q_3 \rangle_{\bar{1}} \geq -1, \quad (2.50)$$

where $\langle \cdot \rangle_{\bar{i}}$ denotes the expectation value when Q_i is not measured.

An observed violation of the inequality is hoped to rule out MR. However, one runs into a major obstacle known as the “clumsiness loophole” [103], which is the need to believe that the measurement would not be disturbing even if MR were true. There have been some attempts to address this, including a scheme set up such that a violation under MR would require two apparently non-disturbing measurements to “collude” in order to result in an overall disturbance [105]. Another option is to use “null result measurements”, where the measurement device interacts with the system only when, say $Q = +1$, but not when $Q = -1$, post-selecting the latter cases [102].

Unfortunately, these options never fully remove the loophole, as they still require some additional model-dependent assumptions. A recent analysis [106] attempts to rigorously pin down exactly what an LGI violation rules out. This is phrased in a

fully model-independent way in terms of operational elements: state preparations, transformations and measurements. It is found that LGIs in fact only test a weaker hypothesis than MR, named “operational eigenstate mixture MR”. An operational eigenstate is essentially a state (or preparation) that always results in a definite value for some Q . The hypothesis then says that the only accessible states are classical mixtures of operational eigenstates. What this means in practice is shown in a recent experiment with superconducting qubits [107]. One first verifies that an intermediate measurement Q_2 of operational eigenstates of Q_1 does not affect the measurement outcomes of Q_3 . Then the usual LGI experiment is performed; the hypothesis is ruled out when a Q_2 measurement is found to disturb the final Q_3 measurement.

In our view, this significantly reduces the utility of LGIs, as the need for additional assumptions weakens what they can demonstrate in a model-independent way.

2.5.2 Collapse models

Alternatively, experimental tests of the superposition principle at macroscopic scales can also be viewed as tests of theories which modify quantum mechanics [53]. Typically, such models predict a rate of “classicalisation” which grows with the difference between branches of a superposition. Then a macroscopic superposition is so short-lived that it is practically never seen. The basis in which decoherence happens differs between models: for example, it could be in the position [52], momentum [108] or energy eigenbasis [109]. In addition, different models predict different scaling behaviour with respect to the size of a superposition in the preferred basis.

A measure of macroscopicity can be associated with each *experiment*, according to how well its demonstration of a coherence time for a certain size of system enables one to rule out parameters in the model [54, 56]. The more macroscopically quantum a state is, and the longer it is maintained, the better it tests a model. While a measure of this type is appealing in its direct connection to experimental physics, it has the disadvantage of being tied to a specific model.

2.5.3 Quantum properties

Finally, a large number of works on macroscopicity stay within the quantum formalism. The general aim is to identify some characteristic quantum feature, and find a

measure which captures the degree to which this is evident on a macroscopic scale. Often, this works by examining a simple model of a Schrödinger’s cat state, such as the N -qubit **Greenberger-Horne-Zeilinger (GHZ) state** [110],

$$\frac{1}{\sqrt{2}}(|0\rangle^{\otimes N} + |1\rangle^{\otimes N}), \quad (2.51)$$

and identifying its characteristic features. We give a brief overview of the past literature in this area.

A number of measures assume that a given state is a superposition of the form $(|\psi_0\rangle + |\psi_1\rangle)/\sqrt{2}$, where there is usually an implicit assumption that the two branches are individually not macroscopically quantum, and try to quantify the extent to which they are macroscopically distinct. Korsbakken et al. [111] ask how many qubits must be measured in order to distinguish between the branches with high probability. In a similar spirit, Marquardt et al. [68] count the average number of operations on single subsystems needed to transform $|\psi_0\rangle$ into $|\psi_1\rangle$. Björk and Mana [66] quantify the relative improvement in the utility of the superposition, over a single branch, for interferometry. Sekatski et al. [112] instead quantify the distinctness of the branches by the resolution required to distinguish them by a measurement of some macroscopic observable. The main problem with such measures is that they only apply to a very specific form of state; in Chapter 6 we show how the latter measure can be extended.

Other proposals are able to treat arbitrary states. Leggett’s “disconnectivity” [65] is one such attempt, aiming to count the number of quantum-correlated particles in a system. Shimizu and Morimae [67], restricting to pure states, suggested that the variance of a macroscopic observable could be used as a measure of macroscopic quantum fluctuations. This was later extended to mixed states by Fröwis and Dür using the quantum Fisher information. For CV systems, Lee and Jeong [70] proposed a quantity that measures non-classicality by the frequency of oscillations in the Wigner function – this was later extended by a related mathematical construction to systems of qubits [113]. Apart from Leggett’s disconnectivity (which is rather tricky to calculate, and has not been well studied), we discuss these measures in detail in Chapters 5 and 7.

Chapter 3

Free operations for speakable coherence

In order to discuss the relevance of coherence for a wide range of quantum information processing tasks, we have to be able to account for situations involving the type of information which has been termed “speakable” [114,115]. In these cases, information is considered fungible, meaning that one is not concerned about the physical medium in which it is stored. In other words, when expressing a state in terms of a basis $|0\rangle, |1\rangle$, one considers “0” and “1” to merely be abstract labels. This is the view taken in classical information theory, and is adopted in very many descriptions of quantum computation and communication.

The seminal work discussing speakable coherence in a resource-theoretic framework is by Baumgratz *et al.* [40]. Similar ideas were presented earlier by Åberg [116], but without formulating a full resource theory. The crucial elements to be pinned down in that work are the free states and free operations. Rather like determining the relevant partition in entanglement theory, a preferred set of basis states must be given at the start. The free states are easily motivated as probabilistic mixtures of the preferred basis states, and are known as **incoherent states** in all subsequent literature. The free operations are correspondingly named **incoherent operations**, and are defined such that they can never create coherence from an incoherent state. There have been numerous following works using this framework to study operational aspects of coherence involving state transformations and communication protocols [117–120,120–122]. Measures derived from this theory have begun to be applied in, for example, state discrimination problems [123], coherent transport phenomena [124,125] and phase transitions in condensed matter systems [126–128].

However, this particular choice of free operations has come under scrutiny. Apart from the work on which this chapter is based, a number of different authors have proposed alternative definitions [41, 129–132]. We shall refer to all classes generically as incoherent operations, naming the original proposal in [40] after the authors as **BCP operations**. With such a variety of approaches, it is clear that an examination of the physical principles and operational interpretations of incoherent operations is necessary in order to arrive at a well-motivated resource theory of speakable coherence.

Here, we define a set of operations within the BCP class with an extra operational restriction, which we refer to as the inability to use coherence. We assume that the required outputs are measurements in the incoherent basis; these operations are then such that coherence of the initial state has no effect on the outputs.

Firstly, we find a mathematical characterisation of these operations, showing them to be equivalent to the **strictly incoherent (SI)** operations which were originally defined in [117]. Next, we prove a dilation theorem showing how the SI operations result from a set of basic incoherent processes including interactions with an ancillary system, and argue that this provides a strong reason for considering the SI set to give a physically justified resource theory of coherence. We also demonstrate the meaning of this dilation picture in terms of interferometry. After giving a new set of measures within this resource theory, we explore relations with the collection of other proposals for speakable coherence.

3.1 Incoherent operations – first attempt

We start by reviewing the earlier proposals for incoherent operations. We work with a d -dimensional Hilbert space $\mathcal{H}$, where d is finite (although many of the concepts can be generalised to infinite dimensions), and specify an orthonormal basis $\{|i\rangle \mid i = 0, \dots, d-1\}$ which we call the **incoherent basis**. Common to all approaches is the assertion that the set $\mathcal{I}$ of free states, or incoherent states, is given by probabilistic mixtures of the incoherent basis states:

$$\mathcal{I} = \left\{ \sum_{i=0}^{d-1} p_i |i\rangle\langle i| \mid p_i \geq 0, \sum_{i=0}^{d-1} p_i = 1 \right\}. \quad (3.1)$$

This sets the most primitive notion of a resource in the theory: we say that every state $\rho \in \mathcal{I}$ has no coherence, while conversely every $\rho \notin \mathcal{I}$ has coherence.¹

Turning to the free operations, note that in general, these could map the Hilbert space $\mathcal{H}$ to a different one $\mathcal{H}'$. However, we focus here on the case $\mathcal{H} = \mathcal{H}'$. In order to arrive at a consistent resource theory, the most fundamental constraint which we must impose on the free operations is that they cannot take a state in $\mathcal{I}$ to one outside $\mathcal{I}$. Let us first examine what this implies for incoherent unitary operations $\rho \rightarrow U\rho U^\dagger$, $U^\dagger U = I$. U must take any incoherent basis state to another one, up to a global phase factor, so that $U|i\rangle = e^{i\phi_i}|f(i)\rangle$, where $f: \{0, \dots, d-1\} \rightarrow \{0, \dots, d-1\}$ and $\phi_i \in \mathbb{R}$. Moreover, it is easy to see that f must be an invertible function, i.e. a permutation. Therefore we can define:

Definition 1. An operator U is an **incoherent unitary** iff it takes the form

$$U = \sum_{i=0}^{d-1} e^{i\phi_i} |f(i)\rangle\langle i|. \quad (3.2)$$

Such an operation is composed of two ingredients: permutations of the incoherent basis states and phase rotations $|i\rangle \rightarrow e^{i\phi_i}|i\rangle$. Both are generally seen as essential elements of the free operations in a resource theory of speakable coherence, so this set of incoherent unitary operations is included in (almost) all proposals.

The situation is more complex for general quantum operations. Working at the level of quantum channels, the weakest constraint can be formulated as follows [116, 117, 132]:

Definition 2. A quantum channel $\mathcal{E}$ is said to be a **maximally incoherent (MI) operation** iff

$$\mathcal{E}(\mathcal{I}) \subseteq \mathcal{I}. \quad (3.3)$$

For a quantum instrument $\{\mathcal{E}_\mu\}$, one can demand that $\mathcal{E}_\mu(\mathcal{I}) \subseteq \mathcal{I}$ for each μ . This ensures that coherence is guaranteed never to be generated. Recall that a channel can generally be associated with many different possible instruments via $\mathcal{E} = \sum_\mu \mathcal{E}_\mu$. The most fine-grained way of doing this is where each element of the instrument is associated with a single Kraus operator, $\mathcal{E}_\mu(\rho) = K_\mu \rho K_\mu^\dagger$. Therefore a stronger condition can be placed on the free operations, as done in [40]:

¹Here we can see another possible generalisation – alternatively, one could specify incoherent subspaces with associated projectors Π_i , satisfying $\Pi_i \Pi_j = \delta_{i,j} \Pi_i$ and $\sum_{i=0}^{d-1} \Pi_i = I$. The set of incoherent states would then be block-diagonal with respect to this partition of $\mathcal{H}$ [116, 117].

Definition 3. A quantum operation $\mathcal{E}$ is said to be a **BCP incoherent** operation iff there exists some Kraus operator decomposition $\mathcal{E}(\rho) = \sum_{\mu} K_{\mu} \rho K_{\mu}^{\dagger}$ such that

$$K_{\mu} \mathcal{I} K_{\mu}^{\dagger} \subseteq \mathcal{I}. \quad (3.4)$$

Convexity of the set $\mathcal{I}$ ensures that condition (3.4) is stronger than (3.3), although it is not immediately clear whether these sets of operations are necessarily different. Winter and Yang [117] demonstrate that indeed $\text{BCP} \subset \text{MI}$ by showing that their respective asymptotic costs are different.

As with unitary operations, it is straightforward to find the form of Kraus operators satisfying (3.4) by looking at the action on incoherent basis states. This motivates the following definition:

Definition 4. K is an **incoherent Kraus operator** iff it satisfies (3.4), or equivalently, iff it is of the form

$$K = \sum_{i=0}^{d-1} c_i |f(i)\rangle \langle i|, \quad (3.5)$$

where $c_i \in \mathbb{C}$ and $f: \{0, \dots, d-1\} \rightarrow \{0, \dots, d-1\}$.

This form is similar to (3.2). It is important to note that, unlike the unitary case, f need not be invertible here.

3.2 Incoherent operations which do not use coherence

Incoherent unitaries and Kraus operators have been defined by the principle that they cannot create coherence. However, we shall see in this section that incoherent unitaries also have another important feature lacking in general incoherent Kraus operators. We term this the inability to use coherence – we shall first motivate and define this.

Consider a situation where a protocol involves transforming a state $\rho \rightarrow \sigma$ and the outputs of interest are determined by measurements in the incoherent basis, i.e. by the probabilities $\langle i|\sigma|i\rangle$. This is relevant in, for instance, transport mechanisms [30, 133], where an excitation travels around a network of sites. If we denote by $|i\rangle$ the state when the excitation is localised at site i , then one is typically interested in the probability $\langle i|\rho(t)|i\rangle$ that it ends up at a certain site i after a given length of

time t . Similarly, in quantum thermodynamics, a natural preferred basis is set by the eigenbasis of the Hamiltonian. Statements about energy transformations (changes in heat and work) can then be made in terms of the probabilities of the system existing in each energy eigenstate [94, 134, 135].

We say that an operation does not use coherence whenever these incoherent outputs are determined only by the incoherent inputs, independent of the coherence. To formalise this, observe that the results of an incoherent measurement on a state ρ are encoded in the “dephased” state $\Phi(\rho) \in \mathcal{I}$ defined by:

Definition 5. The *dephasing operation* is $\Phi: \rho \rightarrow \sum_{i=0}^{d-1} \langle i | \rho | i \rangle | i \rangle \langle i |$.

This is seen to be a channel with $\{|i\rangle\langle i| \mid i = 0, \dots, d-1\}$ as a set of Kraus operators. The result of applying Φ is to remove all coherence from a state while leaving the incoherent measurement probabilities unchanged. One can also think of it as a “measure and forget” operation, where a projective measurement in the incoherent basis is made but a particular outcome is not selected. Therefore, the statement of our principle for an operation with a single Kraus operator becomes:

Definition 6. A Kraus operator K *does not use coherence* iff

$$\Phi(K\rho K^\dagger) = \Phi[K\Phi(\rho)K^\dagger] \quad \forall \rho. \quad (3.6)$$

A characterisation can be found by comparison with condition (3.4) – letting $\mathcal{E}(\rho) = K\rho K^\dagger$, both can be written as

$$(3.4) \Leftrightarrow \mathcal{E} \circ \Phi = \Phi \circ \mathcal{E} \circ \Phi, \quad (3.7)$$

$$(3.6) \Leftrightarrow \Phi \circ \mathcal{E} = \Phi \circ \mathcal{E} \circ \Phi \quad (3.8)$$

Taking the adjoint of the latter, and using $\Phi^\dagger = \Phi$, we have $\mathcal{E}^\dagger \circ \Phi = \Phi \circ \mathcal{E}^\dagger \circ \Phi$. Comparing against (3.4), we see that K satisfies (3.6) iff $K^\dagger$ satisfies (3.7). In this sense, “not creating coherence” and “not using coherence” are dual to each other.

From (3.2), an incoherent unitary has the property

$$U^\dagger = \sum_i e^{-i\phi_i} |i\rangle\langle f(i)| = \sum_j e^{-i\phi_{f^{-1}(j)}} |f^{-1}(j)\rangle\langle j|, \quad (3.9)$$

since f is invertible, so $U^\dagger$ is also incoherent. Thus, for unitaries, Definitions 3 and 6 are equivalent. We can now ask for the general case satisfying both conditions. This

is when K and $K^\dagger$ are both incoherent – corresponding to (3.4) with the restriction that f is invertible. Interestingly, this form of Kraus operator was already defined in Appendix C of [117]; hence, we adopt their naming.

Result 1. *A Kraus operator K neither creates nor uses coherence iff it is of the form*

$$K = \sum_{i=0}^{d-1} c_i |f(i)\rangle\langle i|, \quad (3.10)$$

where $c_i \in \mathbb{C}$ and $f: \{0, \dots, d-1\} \rightarrow \{0, \dots, d-1\}$ is invertible. Following [117], we call K **strictly incoherent (SI)**.

A concise way of expressing the condition for K to be SI can be derived by setting (3.7) to equal (3.8) – this gives $[\mathcal{E}, \Phi] = 0$, where $\mathcal{E}(\rho) = K\rho K^\dagger$.

As with the BCP incoherent operations, we can upgrade this to the level of quantum operations:

Definition 7. *A quantum operation is strictly incoherent (SI) – it neither creates nor uses coherence – iff it has a set of SI Kraus operators.*

By virtue of the corresponding property for its SI Kraus operators, an SI operation $\mathcal{E}$ must satisfy $[\mathcal{E}, \Phi] = 0$. By looking at the conditions (3.7) and (3.8), such an operation can be seen to map diagonal density matrix elements (in the incoherent basis) to diagonals and off-diagonals to off-diagonals. The transformation induced on the diagonal elements $\langle i|\rho|i\rangle \rightarrow \langle i|\mathcal{E}(\rho)|i\rangle$ is characterised by the following:

Result 2. *Under an SI operation, the probabilities $p_i = \langle i|\rho|i\rangle$ transform as*

$$p_i \rightarrow p'_i = \sum_j T_{i,j} p_j, \quad (3.11)$$

where $T_{i,j}$ is a substochastic matrix, meaning that $T_{i,j} \geq 0$ and $\sum_j T_{i,j} \leq 1 \forall i$.

In addition, for any substochastic $T_{i,j}$, there exists an SI operation inducing the transformation (3.11). For an SI channel, replace “substochastic” with “stochastic”, so that $\sum_j T_{i,j} = 1 \forall i$.

Proof. Suppose that the operation is SI, so it has a set of Kraus operators of the form $K_\mu = \sum_i c_i^\mu |f_\mu(i)\rangle\langle i|$, where each f_μ is invertible. From this, we have

$$T_{i,j} = \sum_\mu |c_i^\mu|^2 \delta_{i,f_\mu(j)} \quad (3.12)$$

and so $\sum_j T_{i,j} = \sum_\mu |c_i^\mu|^2$. The Kraus operators must satisfy $\sum_\mu K_\mu^\dagger K_\mu \leq I$, which can be written as

$$\sum_\mu K_\mu^\dagger K_\mu = \sum_{\mu,i,j} (c_i^\mu)^* c_j^\mu |i\rangle \langle f_\mu(i) | f_\mu(j) \rangle \langle j| \quad (3.13)$$

$$= \sum_{\mu,i} |c_i^\mu|^2 |i\rangle \langle i| \stackrel{!}{\leq} I, \quad (3.14)$$

equivalent to $\sum_\mu |c_i^\mu|^2 \leq 1 \forall i$. This proves the first statement.

For the second statement, we define the diagonal matrix $S = \text{diag}(s_0, \dots, s_{d-1})$, where $s_j = \sum_i T_{i,j}$. This has a pseudo-inverse S^{-1} where only the non-zero diagonals are inverted. Observe that the matrix $T' := TS^{-1}$ has row sums $\sum_i T_{i,j}$ equal to 0 or 1. It therefore satisfies the condition of being doubly substochastic, i.e. both T' and T'^T are substochastic.

Theorem 3.2.6 of [136] states that every doubly substochastic matrix can be written as a finite convex combination of partial permutation matrices – that is, $T' = \sum_\mu p_\mu P_\mu$, where each P_μ is a matrix such that every row and column contains at most a single 1 and otherwise all zeros (essentially a permutation matrix with deleted elements). Hence, we have $T = \sum_\mu p_\mu P_\mu S$ and $T_{i,j} = \sum_\mu p_\mu (P_\mu)_{i,j} s_j$, where each $(P_\mu)_{i,j}$ is 0 or 1. Associating P_μ with the permutation f_μ , it may then be seen that the matrix T can be constructed by using the SI Kraus operators

$$K_\mu = \sum_i \sqrt{p_\mu (P_\mu)_{i,f_\mu(i)} s_{f_\mu(i)}} |f_\mu(i)\rangle \langle i|. \quad (3.15)$$

□

As depicted in Fig. 3.1, it is interesting to observe that an SI operation contains an effectively classical stochastic process inside a genuine quantum process. In the next section we describe the physical meaning of SI operations and how they differ from BCP incoherent operations.

3.3 Physical meaning of SI operations

3.3.1 Operations in an interferometer

In this section, we show that the SI operations have a simple physical interpretation, and that it is natural to take both conditions (3.4) and (3.6) for free operations. A full formal discussion is presented in the next section.

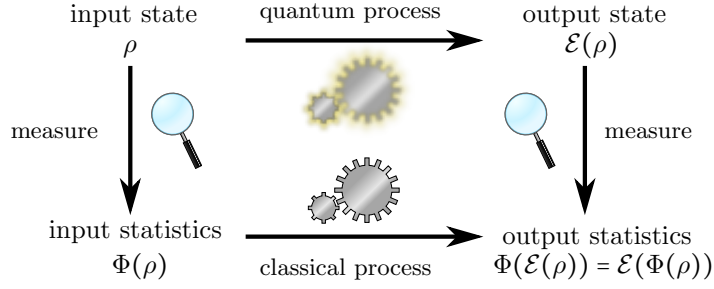


Figure 3.1: Incoherent operations cannot create coherence; strictly incoherent operations in addition cannot use coherence. We depict here an SI operation $\mathcal{E}$, showing that all the initial coherence in the input ρ can be removed with a measurement Φ (in the incoherent basis) without affecting the final measurement outcomes.

Consider an interferometer with a finite number of branches, and suppose that a particle with some internal degrees of freedom passes through it. Writing the Hilbert spaces for the branch and internal degrees of freedom as $\mathcal{H}_B$ and $\mathcal{H}_I$ respectively, the total Hilbert space is $\mathcal{H} = \mathcal{H}_B \otimes \mathcal{H}_I$. The resource for interferometry is coherence between branches, so we choose the incoherent basis to be the set of $|i\rangle_B$ denoting localisation in a branch i . The SI operations then correspond to any combination of path-dependent unitaries on $\mathcal{H}_I$, projective measurements on $\mathcal{H}_I$, and permutations of the paths, followed by tracing out I . This fact will be proved in the next section. See Fig. 3.2 for an illustration.

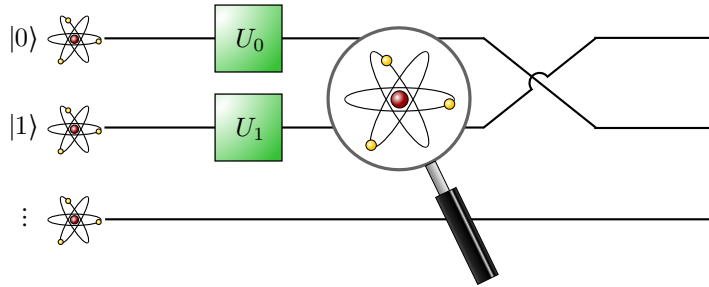


Figure 3.2: Strictly incoherent operations viewed as operations on a particle travelling through an interferometer. The different branches $|i\rangle$ span the Hilbert space $\mathcal{H}_B$, while the internal state of the particle has Hilbert space $\mathcal{H}_I$. The operations are built up by combining path-dependent unitary operations U_i on and measurements of the internal state of the particle, and permutations of the paths.

A typical interferometer is set up with an initial beam splitter, some path-dependent gates encoding phase information, a second beam splitter, and a which-path measurement. The first beam splitter – not considered a free operation here – creates

coherence between branches, while the second enables an informative measurement revealing something about the intermediate phase encoding. While the first beam splitter creates the necessary resource, the second is also crucial, as a direct which-path measurement would give no information at all.

For example, on two branches, a beam splitter followed by a which-path measurement would in combination generate the operation with Kraus operators (acting on $\mathcal{H}_B$)

$$K_0 = |0\rangle\langle +|, K_1 = |1\rangle\langle -|, \quad (3.16)$$

where $|\pm\rangle = (|0\rangle \pm |1\rangle)/\sqrt{2}$. Thus the beam splitter enables a measurement in the coherent $\{|\pm\rangle\}$ basis. This is a BCP incoherent operation, but not SI, as SI operations allow measurements to be only in the incoherent basis.

The lesson is that, within the BCP framework, some resource costs may be hidden within the free operations. Treating the SI operations as the free ones, preparation of coherent states and measurements in a coherent basis are treated on an equal footing.

3.3.2 Dilation

Recall that every quantum operation on a system S can be constructed by taking an initially uncorrelated ancilla α , interacting both parts with a unitary operation, and measuring the ancilla. More precisely, given a set of Kraus operators $\{K_\mu\}$, there exists an initial state $|0\rangle_\alpha$, a global unitary U , and an orthonormal basis $\{|\phi_\mu\rangle_\alpha\}$ such that

$$K_\mu \rho K_\mu^\dagger = \langle \phi_\mu |_\alpha U (\rho_S \otimes |0\rangle\langle 0|_\alpha) U^\dagger | \phi_\mu \rangle_\alpha \quad \forall \mu, \rho \quad (3.17)$$

While this holds for a BCP incoherent operation, we cannot say anything *a priori* about the nature of the corresponding U . Here, we motivate a dilation picture and show that it recovers exactly the SI operations.

We assume that the ancilla α is an additional quantum system for which full control is possible – so restrictions are only placed upon what can be done to the system S . We want to constrain the interaction U such that coherence can never be created on S , even with subsequent operations on α . Suppose the initial state is a pure product state $|i\rangle_S |\psi\rangle_\alpha$. After applying a general U , the resulting state can be written as

$$U|i\rangle_S |\psi\rangle_\alpha = \sum_{j=0}^{d-1} c_j |j\rangle_S |\psi'_j\rangle_\alpha, \quad (3.18)$$

where the $|\psi'_j\rangle$ are normalised but not necessarily orthogonal. We wish to guarantee that no measurement on α can project out a coherent state on S – the following Lemma shows when this is the case.

Lemma 1. *For a pure state $|\Psi\rangle_{S\alpha}$, the projected state $\langle\phi|_\alpha|\Psi\rangle_{S\alpha}$ is incoherent for every $|\phi\rangle_\alpha$ if and only if we can write $|\Psi\rangle_{S\alpha} = |j\rangle_S|\psi'_j\rangle_\alpha$.*

Proof. If $|\Psi\rangle_{S\alpha}$ is of the desired form, then it is clear that the projected state is always proportional to $|j\rangle_S$. Conversely, suppose that $|\Psi\rangle_{S\alpha} = \sum_j c_j |j\rangle_S |\psi'_j\rangle_\alpha$, where at least two distinct c_j are non-zero – without loss of generality, we take these to be c_0 and c_1 . Now define the (unnormalised) state

$$|\phi\rangle = |\psi'_0\rangle + e^{i\theta} |\psi'_1\rangle. \quad (3.19)$$

We have $\langle\psi'_0|\phi\rangle = 1 + e^{i\theta}\langle\psi'_0|\psi'_1\rangle$. It is easy to see that there always exists $\theta \in \mathbb{R}$ such that this inner product is non-zero. Since $\langle\psi'_1|\phi\rangle = e^{i\theta} + \langle\psi'_1|\psi'_0\rangle = e^{2i\theta}\langle\phi|\psi'_0\rangle$, we also have $\langle\psi'_1|\phi\rangle \neq 0$. It follows that

$$\langle\phi|_\alpha|\Psi\rangle_{S\alpha} = c_0\langle\phi|\psi'_0\rangle|0\rangle + c_1\langle\phi|\psi'_1\rangle|1\rangle + \dots, \quad (3.20)$$

where the first two terms are guaranteed to be non-zero, so this state is coherent. $\square$

Applying this, we find that the unitary must have the action $U|i\rangle_S|\psi\rangle_\alpha = |f(i)\rangle_S|\psi'_{f(i)}\rangle_\alpha$ for some function f . It follows that

$$U = \sum_{i=0}^{d-1} |f(i)\rangle\langle i|_S \otimes (U_i)_\alpha, \quad (3.21)$$

and it can be verified that f must be invertible and $U_i^\dagger U_i = I$ in order for $U^\dagger U = I$.

Apart from this interaction, we allow measurements on α – whose outcomes can be recorded and conditioned upon – as well as incoherent unitary operations on S . Note that the function f in (3.21) can always be absorbed into an incoherent unitary on S . Our main result is that these ingredients are necessary and sufficient for performing an SI operation on S .

Result 3. *An operation on a system S is SI iff it can be constructed from the following elementary operations involving an ancilla α :*

1. *Unitary operations on α controlled by the incoherent basis of S : $\sum_i |i\rangle\langle i|_S \otimes (U_i)_\alpha$;*

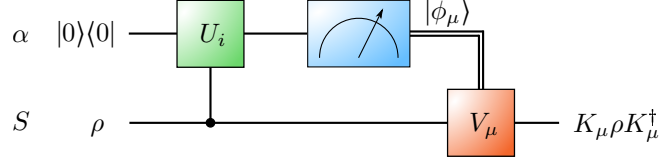


Figure 3.3: Strictly incoherent operations on system S constructed by an incoherent unitary interaction with an ancilla α . An operation is split into a controlled unitary $\sum_i |i\rangle\langle i|_S \otimes (U_i)_\alpha$ followed by a measurement on the ancilla in a basis $|\phi_\mu\rangle$ and an incoherent unitary V_μ on the system conditioned on the measurement outcome μ .

2. *Projective measurements on α in any basis, where selection of outcomes is allowed;*
3. *Incoherent unitary operations on S : $\sum_i e^{i\theta_i} |f(i)\rangle\langle i|$.*

Proof. Firstly, we show that these elementary operations always result in an SI operation. Suppose a protocol is formed by applying an arbitrary number of these in any order. To simplify this, note the following:

- a. Consecutive incoherent unitaries $V = \sum_i e^{i\theta_i} |f(i)\rangle\langle i|$ can be grouped together, as can consecutive controlled-unitaries $U = \sum_i |i\rangle\langle i|_S \otimes (U_i)_\alpha$.
- b. $VU = U'V$, where $U' = \sum_i |i\rangle\langle i|_S \otimes (U_{f^{-1}(i)})_\alpha$. Therefore we can reorder these processes without loss of generality.
- c. Incoherent unitaries on S commute with measurements on α .

As a result, any such protocol can be reduced to a concatenation of sub-processes (of the type shown in Fig. 3.3) consisting of a controlled unitary U followed by a measurement, and an incoherent unitary V at the end – each operation may depend on the results of all previous measurements. Hence, we need only check each sub-process (it is clear the final incoherent unitary is SI).

Suppose a previous measurement has left the ancilla in a state $|\phi_\mu\rangle_\alpha$. Applying a controlled unitary U before projecting onto a state $|\phi_\nu\rangle_\alpha$ corresponds to the Kraus operator

$$K_{\nu|\mu} = \sum_{i=0}^{d-1} \langle \phi_\nu | U_i | \phi_\mu \rangle |i\rangle\langle i|. \quad (3.22)$$

This is SI by inspection, and clearly a combination of these is also SI.

For the converse, suppose an SI operation has a set of Kraus operators $\{K_\mu = \sum_i c_i^\mu |f_\mu(i)\rangle\langle i| \mid \mu = 0, \dots, k-1\}$. We construct an appropriate controlled unitary, ancilla measurement and set of incoherent unitaries to recover this set.

As shown in the proof of Result 2, the condition for an SI operation to be trace-non-increasing is equivalent to $\sum_\mu |c_i^\mu|^2 \leq 1 \ \forall i$. We can always assume that $\sum_\mu K_\mu^\dagger K_\mu = I$ by including another $K_k = \sum_i c_i^k |i\rangle\langle i|$ (which may be zero if the equality is already satisfied).

Since we now have $k+1$ Kraus operators, we choose an ancilla of dimension $k+1$. Given an arbitrary basis set $\{|\phi_\mu\rangle \mid \mu = 0, \dots, k\}$ and an arbitrary initial state $|0\rangle$ for α , we let U_i be such that $U_i |0\rangle = \sum_i c_i^\mu |\phi_\mu\rangle$. After applying U and measuring $|\phi_\mu\rangle$, we conditionally apply $V_\mu = \sum_i |f_\mu(i)\rangle\langle i|$ on S . (See Fig. 3.3 for a circuit representation.) This is seen to reproduce the Kraus operator K_μ . $\square$

This lends weight to the intuitive arguments in the previous section: any operation outside the SI set must involve a global process capable of generating coherence, even if the action on the system happens to be BCP incoherent. Relating back to the interferometer picture, the branch degree of freedom maps onto the system S , and the internal state of the particle acts as the ancilla α . The controlled unitary encapsulates the path-dependent unitaries U_i , while permutations of the branches are represented by incoherent unitaries.

3.4 Monotones

As described in the introduction, in any resource theory, measures of a resource are determined by certain criteria expressed in terms of free states and operations. Baumgratz *et al.* [40] first stated these for coherence; we do the same here, replacing BCP incoherent operations with SI operations.

A measure $C(\rho)$ must be a non-negative real-valued function of the state. The first condition states that a good measure should respect our choice of zero versus non-zero coherence:

$$(C1) \ C(\rho) = 0 \text{ iff } \rho \in \mathcal{I}$$

The condition for monotonicity under free operations can be given in two different ways, corresponding to quantum channels and instruments, respectively. The first

states that coherence cannot increase in a deterministic free process; the second that coherence cannot increase *on average* whenever there is an ensemble of probabilistic outcomes. (In some works the latter version is called “strong monotonicity”.)

(C2a) C is a **deterministic monotone** under any SI channel $\mathcal{E}$:

$$C(\mathcal{E}(\rho)) \leq C(\rho) \quad \forall \rho. \quad (3.23)$$

(C2b) C is an **ensemble monotone** under any SI instrument $\{\mathcal{E}_\mu\}$:

$$\sum_{\mu} p_{\mu} C(\rho_{\mu}) \leq C(\rho) \quad \forall \rho \quad \text{where } p_{\mu} = \text{Tr}[\mathcal{E}_{\mu}(\rho)], p_{\mu} \rho_{\mu} = \mathcal{E}_{\mu}(\rho). \quad (3.24)$$

These are typically considered the most important criteria. There are others which may be seen as desirable, but are not always satisfied for measures of interest. For example:

(C2c) C is a monotone under an SI instrument with a **classical flag**:

$$C\left(\sum_{\mu} p_{\mu} |\mu\rangle\langle\mu| \otimes \mathcal{E}_{\mu}\right) \leq C(\rho) \quad \forall \rho. \quad (3.25)$$

(C3) C is **convex**: for any set of states and probabilities $\{\rho_{\mu}, p_{\mu}\}$ with $\sum_{\mu} p_{\mu} = 1$, we have

$$C\left(\sum_{\mu} p_{\mu} \rho_{\mu}\right) \leq \sum_{\mu} p_{\mu} C(\rho_{\mu}). \quad (3.26)$$

Note that (C2b) and (C2c) are also often typically stated in terms of the most fine-grained instrument for a given channel, i.e. taking a single Kraus operator $\mathcal{E}_{\mu}(\rho) = K_{\mu} \rho K_{\mu}^{\dagger}$. In addition, these conditions are not all independent – for instance, (C2b) and (C3) together imply (C2a).

A number of valid measures are given in [40]. A significant one is the **relative entropy of coherence**:

$$C_R(\rho) := \min_{\sigma \in \mathcal{I}} S(\rho \| \sigma) \quad (3.27)$$

$$= S(\Phi(\rho)) - S(\rho), \quad (3.28)$$

where the second line shows that $C_R(\rho)$ can be expressed as the difference between the von Neumann entropy of ρ and the classical (Shannon) entropy of its diagonal

entries (the minimising incoherent state is $\sigma = \Phi(\rho)$). This satisfies all of the above conditions (C1-3).

A large class of coherence measures can be defined in terms of distance measures. For any contractive pseudo-distance D , we can define the associated coherence measure [40]

$$C_D(\rho) := \min_{\sigma \in \mathcal{I}} D(\rho, \sigma). \quad (3.29)$$

Straightforwardly from the definition, this satisfies (C1) and (C2a); (C3) also holds if D is jointly convex, i.e. if $D(\sum_{\mu} p_{\mu} \rho_{\mu}, \sum_{\mu} p_{\mu} \sigma_{\mu}) \leq \sum_{\mu} p_{\mu} D(\rho_{\mu}, \sigma_{\mu})$.

To assess whether the resource theories defined by BCP and SI operations are different, we can ask whether they have the same set of monotones. SI operations are a strict subset of the BCP operations, so any quantity which is a monotone under BCP operations is automatically a monotone under SI operations. However, it is possible that there exist quantities with monotonicity only under SI operations.

We present a class of distance-based measures which are monotonic under SI operations (using the same assumptions on D as for the measure C_D):

$$C'_D(\rho) := D(\rho, \Phi(\rho)). \quad (3.30)$$

It is clear that this satisfies (C1); in addition, (C2a) follows from the fact that an SI operation $\mathcal{E}$ is such that $[\mathcal{E}, \Phi] = 0$:

$$\begin{aligned} C'_D(\mathcal{E}(\rho)) &= D(\mathcal{E}(\rho), \Phi \circ \mathcal{E}(\rho)) \\ &= D(\mathcal{E}(\rho), \mathcal{E} \circ \Phi(\rho)) \\ &\leq D(\rho, \Phi(\rho)) \\ &\leq C'_D(\rho). \end{aligned} \quad (3.31)$$

In some cases, we may find $C'_D = C_D$. For example, this is true for the measure derived from the distance using the l_1 norm, $D_{l_1}(\rho, \sigma) = \|\rho - \sigma\|_{l_1} = \sum_{i,j} |\rho_{i,j} - \sigma_{i,j}|$ [40] (where $\rho_{i,j} = \langle i | \rho | j \rangle$):

$$C_{l_1}(\rho) := \min_{\sigma \in \mathcal{I}} D_{l_1}(\rho, \sigma) = D_{l_1}(\rho, \Phi(\rho)) = \sum_{i \neq j} |\rho_{i,j}|. \quad (3.32)$$

On the other hand, the measures based on the fidelity $C_F(\rho) := \min_{\sigma \in \mathcal{I}} [1 - F(\rho, \sigma)^2]$ [40, 137, 138] and the trace norm $C_{tr}(\rho) := \min_{\sigma \in \mathcal{I}} \|\rho - \sigma\|_{tr}$ [40, 139] do not generally have $\Phi(\rho)$ as the closest incoherent state. (Note that some variations upon C_F exist

in the literature, including $\min_{\sigma \in \mathcal{I}}[1 - F(\rho, \sigma)]$, which fails (C2b) [137].) Hence C'_F and C'_{tr} are SI monotones but not necessarily BCP monotones. However, we have been unable to find examples where these measures are able to increase under BCP operations. As no proof of monotonicity exists so far, it remains an open question whether these are also monotonic under BCP, and therefore whether SI operations admit a strictly larger set of monotones.

One might want to search for examples of BCP operations where C'_F fails (C2a) – however, note that this can never happen for $d = 2$ dimensions. Chitambar and Gour [130] have shown that, for qubits, any deterministic transformation $\rho \rightarrow \sigma$ which can be done by BCP operations can also be done by SI operations. It follows that any deterministic monotone under SI operations is a deterministic monotone under BCP operations in this case.

3.5 Comparison with other incoherent operations

In this section, we examine the relation between SI operations and other classes of incoherent operations which have been proposed. These are all subsets of the previously mentioned MI operations at the top of the hierarchy.

3.5.1 Translationally invariant operations

Given an observable $A = \sum_i a_i |i\rangle\langle i|$ whose eigenbasis is chosen to coincide with the incoherent basis, we have the following [41, 115]:

Definition 8. A quantum operation $\mathcal{E}$ is **translationally invariant (TI)** iff

$$e^{-itA} \mathcal{E}(\rho) e^{itA} = \mathcal{E}(e^{-itA} \rho e^{itA}) \quad \forall t \in \mathbb{R}, \rho. \quad (3.33)$$

Equivalently, one may say that $\mathcal{E}$ is **covariant** with respect to the unitary transformations $\mathcal{U}_t(\rho) = e^{-itA} \rho e^{itA}$ generated by A . This can also be expressed as $[\mathcal{E}, \mathcal{U}_t] = 0 \quad \forall t$.

As will be discussed in detail in Chapter 5, TI operations strictly speaking give rise to a resource theory of unspeakable coherence; however, we include them here since the comparison is illuminating. We have chosen here not to allow incoherent subspaces for speakable coherence, so we similarly assume that the spectrum of A is non-degenerate. Unlike BCP and SI operations, the definition of TI operations at the

level of total quantum channels actually implies the existence of a set of incoherent Kraus operators of the form [101]

$$K_\delta = \sum_{i,j: a_i - a_j = \delta} c_{i,j} |i\rangle\langle j|. \quad (3.34)$$

Comparing with (3.10), we see that these Kraus operators are in the SI set, and are restricted such that the function f maps i to j satisfying $a_j = a_i + \delta$, where δ is a constant for each Kraus operator. Because A is non-degenerate, there is at most one such value of j for any given i . An arbitrary reshuffling of basis states is not allowed – each Kraus operator may only rigidly shift states up or down in the spectrum of A . Coherence associated with off-diagonal elements $|i\rangle\langle j|$ for different values of $a_i - a_j$ constitute separate resources which cannot be interconverted under TI operations. This makes TI operations a strict subset of SI. They do not include the full set of incoherent unitaries that we require for a speakable coherence resource theory – only $U = \sum_i e^{i\theta_i} |i\rangle\langle i|$ is possible (it may also be possible to shift states in the spectrum of A unitarily if this spectrum is unbounded from above and below).

The physical motivation for TI operations is very strong, deriving from conservation laws and considerations about reference frames – see Chapter 5. Importantly, they have a dilation theorem [41, 101, 140]. This invokes an ancilla α with an associated observable $B = \sum_j b_j |j\rangle\langle j|$ defining its own incoherent basis. Then any TI operation $\mathcal{E}$ can be constructed via

$$\mathcal{E}(\rho_S) = \text{Tr}_\alpha \left[(I_S \otimes E_\alpha) U (\rho_S \otimes \sigma_\alpha) U^\dagger \right], \quad (3.35)$$

where the state σ_α and POVM element E_α are both incoherent (with respect to $\{|j\rangle\}$), and U is covariant with respect to transformations generated by $A \otimes I + I \otimes B$. Physically, U corresponds to an interaction conserving the global additive quantity $A \otimes I + I \otimes B$; no coherent resources are needed to prepare the initial ancilla state or to perform its final measurement.

Unlike the SI dilation, where full control of the ancilla is assumed for free, here the ancilla can be used as a resource in the following sense. The TI set captures all operations which can be performed without access to a reference for the phase conjugate to A . However, a coherent ancilla acts as a phase reference, and thus can be used to overcome the restriction imposed by a covariant interaction U [141–144]. The reason for this difference is that the free interaction U in the SI case is of the

form (3.21), which can never generate a superposition of incoherent basis states (even before tracing out the ancilla). However, for the TI case, the free U can generate such superpositions. Consider, for instance, that A and B are both number observables; then, U is allowed to create a state such as $|0\rangle|1\rangle + |1\rangle|0\rangle$. In the TI resource theory, it is assumed that, when systems are combined, the global observable $A \otimes I + I \otimes B$ is the one which defines incoherent states – and this may have degenerate subspaces in which superpositions are considered free.

A significant subset of TI operations is commonly considered in quantum thermodynamics. Given a Hamiltonian H and a heat bath at temperature $T = 1/(k_B\beta)$, a resource theory can be defined with a single free state: the thermal (or Gibbs) state $\gamma = e^{-\beta H}/\text{Tr}[e^{-\beta H}]$. The set of **thermal operations** is generated by energy-conserving interactions between the system and an ancilla initially in a thermal state [94, 134, 145]. These are a strict subset of TI operations with respect to the energy eigenbasis. This framework therefore allows for the discussion of coherence as resource in quantum thermodynamics [135, 146, 147].

3.5.2 Genuinely and fully incoherent operations

Streltsov [129] proposed a set of incoherent operations as follows:

Definition 9. A quantum channel $\mathcal{E}$ is **genuinely incoherent (GI)** iff it preserves every incoherent state:

$$\mathcal{E}(\rho) = \rho \quad \forall \rho \in \mathcal{I}. \quad (3.36)$$

These have the property that *every* possible set of Kraus operators is diagonal in the incoherent basis: $K_\mu = \sum_i c_i^\mu |i\rangle\langle i|$. As such, the GI set is a strict subset of TI operations. Their power is very limited, since even different incoherent pure states cannot generally be interchanged. The associated dilation is the same as in Result 3, but without the ability to perform incoherent unitaries on the system – GI operations are SI operations without permutations.

A number of results about pure and mixed state transformations under GI operations is given in [131]. There, the authors aim to construct a more powerful set which keeps the property that all Kraus operator representations are incoherent. Therefore they define:

Definition 10. A quantum operation is **fully incoherent (FI)** iff every set of corresponding Kraus operators is incoherent.

Equivalently, they prove that these are such that all Kraus operators are all incoherent *and* of the same form (i.e., have the same pattern of non-zero matrix elements in the incoherent basis). Rather strangely, the FI set is not convex, casting doubt on its physical consistency. This is demonstrated in [131] via the qubit example

$$\mathcal{E}(\rho) = p\sigma^x\rho\sigma^x + (1-p)\sigma^z\rho\sigma^z, \quad (3.37)$$

which is not FI for $p \in (0, 1)$, while each of the two terms is FI. This example also shows that an SI operation is not necessarily FI.

Conversely, it is clear that not every FI operation is SI – consider, for example, the operation with a single Kraus operator of the form $K = a|0\rangle\langle 0| + b|0\rangle\langle 1|$. This fact is stated in [131]; however, a subtlety is missed there that FI *channels* can be a subset of SI channels, which involves imposing $\sum_\mu K_\mu^\dagger K_\mu = I$. We prove that exact relation depends on the dimension:

Result 4. For $d = 2$ or 3 , the FI channels are a subset of the SI channels. For $d \geq 4$, there exist FI channels which are not SI.

Proof. We begin by identifying a useful way of expressing FI Kraus operators. Note firstly that any incoherent Kraus operator can be written as

$$K_\mu = \sum_i |i\rangle\langle\psi_i^\mu|, \quad (3.38)$$

where each $|\psi_i^\mu\rangle$ need not be normalised. In the FI case, the statement that all K_μ are of the same form amounts to the following: Let $\{S_i \mid i = 0, \dots, d-1\}$ be a collection of disjoint (possibly empty) subsets of $\{0, \dots, d-1\}$ such their union covers the whole set. Defining $\mathcal{H}_i = \text{span}\{|j\rangle \mid j \in S_i\}$, we have $|\psi_i^\mu\rangle \in \mathcal{H}_i \forall \mu$.

For an FI channel, $I = \sum_\mu K_\mu^\dagger K_\mu = \sum_{\mu,i,j} |\psi_i^\mu\rangle\langle i|j\rangle\langle\psi_j^\mu| = \sum_{\mu,i} |\psi_i^\mu\rangle\langle\psi_i^\mu|$. Since the $|\psi_i^\mu\rangle$ belong to the subspaces $\mathcal{H}_i$, which are orthogonal for different i , it follows that

$$\sum_\mu |\psi_i^\mu\rangle\langle\psi_i^\mu| = \sum_{j \in S_i} |j\rangle\langle j| \quad \forall i. \quad (3.39)$$

Considering the rank of both sides of this equation, note that it requires the number of μ values to be at least $|S_i|$. We now apply this to cases of specific dimensions.

- $d = 2$: Up to a relabelling of the incoherent basis, the only Kraus operator form, that is FI but not SI, is $K_\mu = |0\rangle\langle\psi^\mu|$. At least two values of μ are needed to satisfy (3.39). There always exists a minimal Kraus representation in which the K_μ are linearly independent (Exercise 6.14 of [148]), and so the $|\psi^\mu\rangle$ are also linearly independent. Therefore two values $\mu = 0, 1$ are sufficient.

We now use Theorem 2.6 of [72], which says that two sets $\{|\alpha_i\rangle\}$, $\{|\beta_j\rangle\}$ generate the same Hermitian operator, i.e. $\sum_i |\alpha_i\rangle\langle\alpha_i| = \sum_j |\beta_j\rangle\langle\beta_j|$, iff there exists a unitary matrix u_{ij} such that $|\alpha_i\rangle = \sum_j u_{ij} |\beta_j\rangle$. (If the sets are different sizes, then we just pad the smaller one with zero entries to equal the length of the other.) Given (3.39), we can therefore write $|\mu\rangle = \sum_\nu u_{\mu\nu} |\psi^\nu\rangle$.

If we now let $L_\mu = |0\rangle\langle\mu|$ ($\mu = 0, 1$), then it can be seen that $L_\mu = \sum_\nu u_{\mu\nu}^* K_\nu$. Two sets of Kraus operators are related by a unitary remixing in this way iff they generate the same channel. This is Theorem 8.2 of [72], and is a consequence of the previously used Theorem (the same comment about the length of the two sets applies).

Hence we have shown that L_μ are a valid set of Kraus operators for the channel, and they are clearly SI.

- $d = 3$: There are now two FI Kraus operator forms that are not SI. The first is

$$K_\mu = \begin{pmatrix} * & * & * \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = |0\rangle\langle\psi^\mu|, \quad (3.40)$$

where $*$ denotes any potentially non-zero element. The logic is exactly as for $d = 2$: we can always assume a channel decomposition with three values $\mu = 0, 1, 2$. As before, this leads to an equivalent set of three SI Kraus operators $L_\mu = |0\rangle\langle\mu|$.

Secondly, there is the form

$$K_\mu = \begin{pmatrix} * & * & 0 \\ 0 & 0 & * \\ 0 & 0 & 0 \end{pmatrix} = |0\rangle\langle\psi^\mu| + c^\mu |1\rangle\langle 2|, \quad |\psi^\mu\rangle \in \text{span}\{|0\rangle, |1\rangle\}. \quad (3.41)$$

We can generally assume three Kraus operators of this form (due to linear independence of the minimal set). Equation (3.39) states that $\sum_{\mu=0}^2 |\psi^\mu\rangle\langle\psi^\mu| = |0\rangle\langle 0| + |1\rangle\langle 1|$. As before, this guarantees the existence of a unitary matrix $u_{\mu\nu}$

such that $|0\rangle = \sum_{\mu} u_{0\mu} |\psi^{\mu}\rangle$, $|1\rangle = \sum_{\mu} u_{1\mu} |\psi^{\mu}\rangle$, $0 = \sum_{\mu} u_{2\mu} |\psi^{\mu}\rangle$. The corresponding new set of Kraus operators $L_{\mu} = \sum_{\nu} u_{\mu\nu}^* K_{\nu}$ can be written as

$$L_0 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & * \\ 0 & 0 & 0 \end{pmatrix}, L_1 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & * \\ 0 & 0 & 0 \end{pmatrix}, L_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & * \\ 0 & 0 & 0 \end{pmatrix}, \quad (3.42)$$

which are SI.

- $d \geq 4$: We give a counter-example for $d = 4$ (which also exists in higher dimensions by embedding). Define a channel with two Kraus operators

$$K_0 = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, K_1 = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \quad (3.43)$$

It is easily verified that these satisfy $K_0^{\dagger} K_0 + K_1^{\dagger} K_1 = I$. There are only two possible linear combinations that leave a single non-zero element in the first row, namely $K_0 \pm K_1$ (up to a constant). However, the second row then contains two non-zero elements. Therefore no linear combination can generate a set of SI Kraus operators.

□

3.5.3 Physical incoherent operations

Chitambar and Gour also address the physical meaning of incoherent operations by considering their dilation [130]. However, they ask for a type of dilation which is more restricted than that in Result 3. Similarly to the TI dilation, the ancilla α also has its own incoherent basis $\{|j\rangle\}$. The global incoherent basis for $\mathcal{H}_S \otimes \mathcal{H}_{\alpha}$ is then defined to be the product basis $\{|i\rangle_S |j\rangle_{\alpha} \mid i = 0, \dots, d_S - 1; j = 0, \dots, d_{\alpha} - 1\}^2$. The interaction U is assumed to be incoherent with respect to this basis; one can then perform incoherent projective measurements on α . The resulting operations are called **physical incoherent (PI) operations**, and always have a set Kraus operators of the form

$$K_{\mu} = \sum_i e^{i\theta_i} |f_{\mu}(i)\rangle \langle i| P_{\mu}, \quad (3.44)$$

where each f_{μ} is invertible and the P_{μ} form a complete and orthogonal set of incoherent projectors.

PI operations are a strict subset of SI operations; moreover, they are rather restricted in the transformations they can perform. For example, a pure state $|\psi\rangle$ can be converted to $|\phi\rangle$ under PI operations with non-zero probability only if $|\psi\rangle$ and $|\phi\rangle$ are connected by an incoherent unitary [130]. This rules out, for instance, $(|0\rangle + |1\rangle)/\sqrt{2} \rightarrow \sqrt{p}|0\rangle + \sqrt{1-p}|1\rangle$ for any $p \neq 1/2$, while this is possible under SI operations [117, 130].

Analogously to Result 2, we note that the transformation on the diagonal elements of a state $p_i \rightarrow p'_i = \sum_j T_{i,j} p_j$ is associated with a *doubly* substochastic matrix $T_{i,j}$. By a simpler version of the previous proof (without using the matrix S), it also follows that every such doubly substochastic transformation can be achieved with a PI operation.

3.5.4 Dephasing-covariant incoherent operations

Looking for a set of operations more powerful than PI, a new set is suggested in [130] – the same set is independently proposed by Marvian and Spekkens [41]. This can be seen as moving Definition 6 about “not using coherence” to the level of operations rather than Kraus operators:

Definition 11. *A quantum operation $\mathcal{E}$ is a **dephasing-covariant incoherent (DI) operation** iff it commutes with dephasing: $[\mathcal{E}, \Phi] = 0$.*

Evidently every SI operation is also DI. We first show that the converse holds for qubits.

Result 5. *For $d = 2$, the sets of SI and DI operations are the same.*

Proof. It has already been shown that $SI \subseteq DI$. For the converse, we use the Choi-Jamiołkowski isomorphism [72] (Chapter 8) which associates every quantum operation on S with a (sub-normalised) state on RS , where R is another system with the same

²It is claimed in [130] that this form for the global incoherent states is fixed by the “no super-activation postulate”: if ρ and σ are both free states for their respective subsystems, then $\rho \otimes \sigma$ is free. This claim is false, as shown by the TI dilation picture. The postulate only determines that the global incoherent states must *contain* $|i\rangle|j\rangle$. In order to rule out the existence of a larger set, we would have to assume a postulate such as the following: all global free states can be obtained from statistical mixtures of states of the form $\rho \otimes \sigma$, where ρ and σ are free. This arises naturally when the set of global free operations is given by local free operations and classical communication.

dimension as S :

$$\sigma := (I_R \otimes \mathcal{E}_S)(|\alpha\rangle\langle\alpha|) \quad (3.45)$$

$$= \sum_{i,j} |i\rangle\langle j|_R \otimes \mathcal{E}_S(|i\rangle\langle j|), \quad (3.46)$$

where $|\alpha\rangle = \sum_i |i\rangle_R |i\rangle_S$. The condition $\sigma \geq 0$ is equivalent to saying that $\mathcal{E}$ is CP. Given that a DI operation must map diagonals to diagonals and off-diagonals to off-diagonals, the most general form for σ in the basis $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$ is

$$\sigma = \begin{pmatrix} a & 0 & 0 & b \\ 0 & c & d & 0 \\ 0 & d^* & e & 0 \\ b^* & 0 & 0 & f \end{pmatrix}, \quad (3.47)$$

where $a, c, e, f \geq 0$ and $b, d \in \mathbb{C}$. This is block-diagonal and hence can be written as $\sigma = \sum_{\mu=0}^3 |\psi_\mu\rangle\langle\psi_\mu|$, where $|\psi_{0,1}\rangle \in \text{span}\{|00\rangle, |11\rangle\}$ and $|\psi_{2,3}\rangle \in \text{span}\{|01\rangle, |10\rangle\}$ (not normalised). We can write $|\psi_0\rangle = x_0 |00\rangle + y_0 |11\rangle$, so

$$|\psi_0\rangle\langle\psi_0| = \begin{pmatrix} |x_0|^2 & 0 & 0 & x_0 y_0^* \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ x_0^* y_0 & 0 & 0 & |y_0|^2 \end{pmatrix}. \quad (3.48)$$

Note that this corresponds to the Choi-Jamiołkowski form of the operation $\rho \rightarrow K_0 \rho K_0^\dagger$, where

$$K_0 = \begin{pmatrix} x_0 & 0 \\ 0 & y_0 \end{pmatrix}. \quad (3.49)$$

Similarly we can associate the remaining $|\psi_\mu\rangle$ with

$$K_1 = \begin{pmatrix} x_1 & 0 \\ 0 & y_1 \end{pmatrix}, K_2 = \begin{pmatrix} 0 & x_2 \\ y_2 & 0 \end{pmatrix}, K_3 = \begin{pmatrix} 0 & x_3 \\ y_3 & 0 \end{pmatrix}. \quad (3.50)$$

Therefore $\mathcal{E}$ has a set of SI Kraus operators. □

The proof method is blocked for higher dimensions because the Choi matrix σ no longer has a block-diagonal form. In fact, we have an example that disproves the converse for $d \geq 4$. Unfortunately, the case of $d = 3$ is left unanswered.

Result 6. *For $d \geq 4$, there exist channels that are DI but not SI.*

Proof. The statement is proved by the FI example (3.43), which we have shown does not represent an SI channel. In [131] it is stated that every FI operation is DI. To be

precise, this is in fact not true at the level of quantum operations (consider a single incoherent Kraus operator) – but is true for channels. We repeat the proof here, using our terminology. We write a set of FI Kraus operators as $K_\mu = \sum_i |i\rangle\langle\psi_i^\mu|$ in the notation of the proof of Result 4. Then we calculate the matrix element

$$\begin{aligned}\langle k|\left(\sum_\mu K_\mu|i\rangle\langle j|K_\mu^\dagger\right)|k\rangle &= \sum_{\mu,l,m}\langle k|l\rangle\langle\psi_l^\mu|i\rangle\langle j|\psi_m^\mu\rangle\langle m|k\rangle \\ &= \sum_\mu\langle\psi_k^\mu|i\rangle\langle j|\psi_k^\mu\rangle \\ &= \langle j|\left(\sum_\mu|\psi_k^\mu\rangle\langle\psi_k^\mu|\right)|i\rangle.\end{aligned}\tag{3.51}$$

Using (3.39), which expresses $\sum_\mu K_\mu^\dagger K_\mu = I$, this vanishes whenever $i \neq j$. Hence the channel is DI. $\square$

In [41], an attempt is made for the physical justification of the DI set via a dilation theorem. They show that, given a dilation with an incoherent ancilla, a dephasing-covariant global unitary, and an incoherent ancilla measurement, the resulting operation is DI. They leave the converse as an open question. In fact, we can see that a dephasing-covariant global unitary U is in general of the same form (3.21) as used in our SI dilation. However, their dilation does not allow for a conditional incoherent unitary V_μ on the system. Therefore, such a dilation necessarily gives an SI operation in which the permutations f_μ performed by all Kraus operators are the same. This is a strict subset of SI, so their dilation cannot recover the full DI set.

3.5.5 Coherent transport and cooling maps

One of the bodies of work calling for rigorously defined measures of coherence is the study of coherent transport processes. Working in this area, Levi and Mintert [133] propose a quantitative measure of coherent delocalisation. Without defining a full resource theory, they justify this measure by proving its monotonicity under a set of incoherent processes. They assume that a single excitation is present; denoting by $|i\rangle$ the state of an excitation localised at site i , their incoherent processes fall into two classes:³

1. “Modification of phase coherence”: Kraus operators $A_i = u_1|i\rangle\langle i| + u_2\sum_{j\neq i}|j\rangle\langle j|$.
2. “Incoherent hopping”: Kraus operators $B_{ji} = |j\rangle\langle i|$.

These clearly form a strict subset of SI operations. The Kraus operators of the form A_i are simply diagonal – it is also evident that, by composing $A_i^{(1)} A_j^{(2)} \dots$ with different values of $u_{1,2}$, the full set of diagonal SI Kraus operators $\sum_i c_i^\mu |i\rangle\langle i|$ can be generated. However, the B_{ji} are too restrictive to recover the full SI set, or even the PI set. Each B_{ji} corresponds physically to a projective measurement onto a single site i , which is then transferred to site j . With general SI operations, it is possible to simultaneously interchange two pairs of sites while retaining coherence. For example, the transformation $|0\rangle + |1\rangle \rightarrow |1\rangle + |2\rangle$ can be done with an SI operation but not with any application of B_{10} or B_{21} .

Almost the same set of operations appears in a different physical context. The authors of [149] show that thermal operations with a low-temperature bath take on a certain form. The Kraus operators are either of the form $K = \sum_i \lambda_j |E_i\rangle\langle E_i|$ or $J_{jk} = \mu_{jk} |E_j\rangle\langle E_k|$, where j and k are such that $E_j < E_k$. The set of “cooling maps” is defined to contain all channels whose Kraus operators can be written in this way. As compared with the above set, the only difference is the order imposed on the allowed J_{jk} .

3.6 Conclusions

In this chapter, we have defined the condition that states when a quantum operation does not use coherence, and have seen that this can be viewed as dual to the condition that an operation does not create coherence. Asking for both conditions to be satisfied simultaneously gives the set of SI operations. These constitute a set of processes living somehow on the boundary between quantum and classical, where coherence can be present but does not contribute to the measured outcomes.

We have shown that SI operations admit a description with basic incoherent unitary gates and measurements involving interactions with an ancilla, and that there is a corresponding intuitive interpretation in terms of operations on an interferometer. We argue that this makes the SI operations the strongest physically motivated set of incoherent operations so far. We have seen that other sets can also be proposed

³In [133], these processes are argued to be determined by two conditions: they preserve the number of excitations, and an initially localised state is kept localised. The first condition simply ensures that we can consistently work in the single-excitation subspace; the second can be recognised as the inability to create coherence. Since we know that the largest such set (determined by Kraus operators) is the BCP incoherent set, the authors’ logic must be flawed.

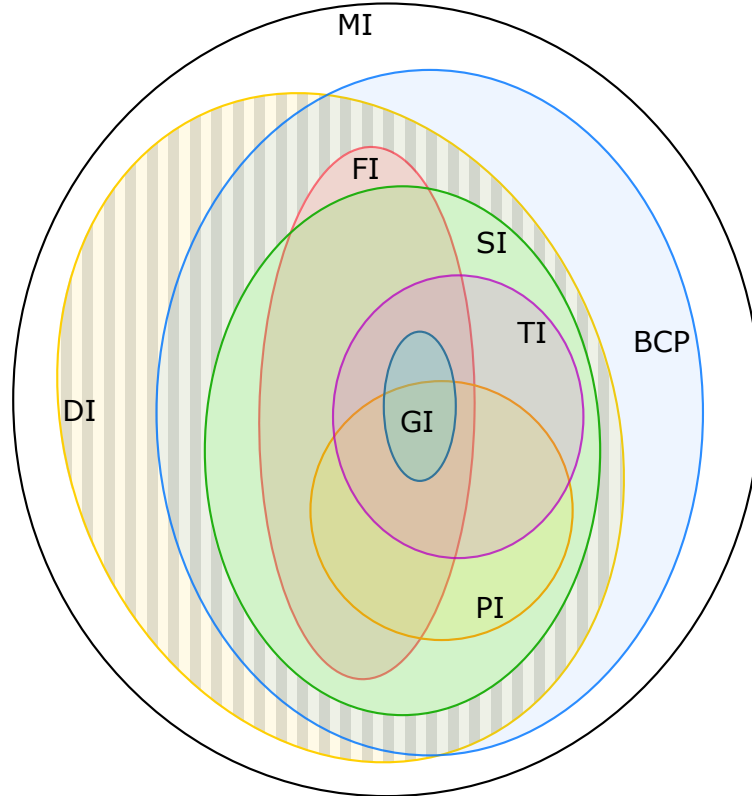


Figure 3.4: Schematic representation of the relations between some sets of incoherent operations, at the level of channels rather than Kraus operators. It is unknown whether the region shaded with bars is empty.

from physical considerations (for example, translationally invariant operations, thermal operations, and incoherent transport processes), but these do not include the full set of incoherent unitaries which are desired of a resource theory of speakable coherence. The set of PI operations, while physically consistent in the sense of having an incoherent dilation, is also too restrictive for a well-behaved theory.

Another proposal in the literature, namely DI operations, is very close to SI, defined as whole maps (without restrictions on individual Kraus operators) which can neither create nor use coherence. However, we have shown that the DI set is in general (for dimension greater than three) strictly larger than SI. Therefore our dilation cannot be applied, and it is not known whether there is some more general dilation giving rise to the DI set. It is at least hard to see how to construct a physically reasonable dilation even “larger” than the one we have given.

There are a number of open resource-theoretic questions pertaining to SI operations. Firstly, while we have found an additional set of SI monotones, it is still

unknown whether or not these are strictly a new set – i.e., whether they are also monotones under the BCP set. An interesting possibility remains that the monotones for the two sets of operations are actually the same. Either outcome would have intriguing implications for the operational differences between BCP and SI operations.

Similarly, one can compare the state transformations possible under both SI and BCP operations. Winter and Yang [117] have shown that the pure state transformations are identical, as are the rates of formation. The only one of their results which has not been extended to SI operations is distillation from mixed states. Their proof of the distillation rate for BCP operations makes explicit and non-trivial use of measurements in a coherent basis. A set of desirable characteristics for a resource theory of speakable coherence is given in [15]. The last of these is “no bound coherence”, meaning that every coherent state should be (asymptotically) distillable at a non-zero rate – which is an open question for SI operations.

Finally, it remains to be seen whether our results hold for the possible generalisations mentioned earlier, namely for incoherent subspaces and for operations which map between different state spaces. In addition, there is much to be discussed about distributed scenarios in which separated observers are restricted to local incoherent operations and classical communication: so far, only BCP operations have been used in this context [118–120].

Chapter 4

Relating coherence and discord

This chapter explores properties of quantum coherence in multipartite systems (composed of two subsystems or more). In settings of this type, one can discuss correlations and the extent to which they are useful resources. Relations between the resource theories of coherence and entanglement have been studied in a number of works [118, 138, 150], but here we focus on another quantum correlation known as quantum discord [44, 151]. As we explain below, this was originally introduced in two independent works [42, 43] which noticed that there is generally a gap between the quantum mutual information, a measure of the degree of correlation between two subsystems, and the classical correlations between local measurement outcomes. The states for which such a gap exists have a quantum-correlated character that includes but goes beyond entanglement.

Discord is a potentially exciting candidate as a quantum resource in part because it is far less fragile [152] than entanglement, which can often vanish under noise in a phenomenon known as “sudden death” [153]. It is known to signal a quantum advantage in a number of protocols including state merging [154, 155], remote state preparation [156], cryptography [157], and the encoding of information in a bipartite state [158, 159]; it also plays a role in quantum locking of classical correlations [160] and non-local effects in thermodynamics [161, 162].

Therefore it is pertinent to ask to what extent coherence and discord are interchangeable – when can one resource be traded for the other? In Section 4.2, we provide some answers by giving bounds on the discord which can be created between subsystems under incoherent operations, in terms of the coherence of the subsystems. We also use this to aid the discussion of coherence and discord as resources for a type

of noisy quantum computation [45, 163].

Despite the operational interpretations which have been found for discord, its general status as a resource lacks a firm footing because there is no known resource theory. The difficulty is that the appropriate set of free operations is unclear; surprisingly, discord can increase under some local operations [164]. We attempt to offer some insight into this problem in Section 4.3.3 by studying the behaviour of discord under the free operations for coherence. To do this, we work with a simpler basis-dependent version of discord. We demonstrate monotonicity properties under the SI operations introduced in Chapter 3, with increases in discord being attributed to the loss of a classical memory. Finally, we find a new connection between free operations for coherence and discord by considering operations that are SI with respect to every basis.

4.1 Basics of quantum discord

As was noted in [42, 43] that while the classical mutual information can be defined in two equivalent ways, these two expressions do not always agree in the quantum case. The quantum **mutual information** between two subsystems A and B is defined as

$$I(A : B)_\rho := S(\rho_A) + S(\rho_B) - S(\rho_{AB}). \quad (4.1)$$

This can be thought of as the amount of information stored in the state that is shared jointly between the subsystems. It vanishes exactly for product states and is maximised for the maximally entangled state. In the case of two classical random variables X and Y , the analogous expression takes the same form but with the Shannon entropy in place of the von Neumann entropy, i.e. $I_{cl}(X : Y) = H(Y) + H(X) - H(X, Y)$. It is simple to show that this may be rewritten as $I_{cl}(X : Y) = H(X) - \sum_x p_x H(Y|X = x)$, where p_x is the marginal distribution of X and $H(Y|X = x)$ is the Shannon entropy of the Y distribution $p_{y|x} = p_{x,y}/p_x$ obtained when the outcome $X = x$ is measured.

In the quantum case, suppose a measurement is made with POVM elements $\{A_a\}$ on A . When the outcome A_a is obtained, the state of B is projected to $\rho_{B|a} = \text{Tr}_A(A_a \rho_{AB})/p_a$, where $p_a = \text{Tr}_{AB}(A_a \rho_{AB})$. Using the same construction as the classical quantity, we can define

$$J(B|\{A_a\})_\rho := S(\rho_B) - \sum_a p_a S(\rho_{B|a}). \quad (4.2)$$

However, this is in general *less* than $I(A : B)_\rho$. Even after maximising over all possible POVMs, the resulting quantity

$$J(B|A)_\rho := \max_{\{A_a\}} J(B|\{A_a\})_\rho \quad (4.3)$$

is still less than $I(A : B)_\rho$. Following [42]:

Definition 12. The **discord** of a bipartite state ρ_{AB} is defined as the gap

$$\delta(B|A)_\rho := I(A : B)_\rho - J(B|A)_\rho. \quad (4.4)$$

The rationale behind this definition is that it neatly expresses the “total correlations” $I(A : B)$ as a sum of a quantum part $\delta(B|A)$ and a classical part $J(B|A)$. Such a splitting is reasonable because $\delta(B|A) \geq 0$; this follows from the fact that the mutual information cannot increase under a local measurement [165].

In [42], the measurement on A was assumed to be a set of rank-one projectors – it was shown later that the minimum value in the definition of $\delta(B|A)$ (corresponding to the maximum in $J(B|A)$) is in fact achieved by rank-one projectors (Theorem 5 of [166]). In some literature, and in the results presented below, the POVM elements may be further restricted to a set of *orthogonal* one-dimensional projectors. In this case, the optimisation does not generally achieve the value given by general measurements [167, 168].

This quantity is generally asymmetric in the sense that $\delta(B|A) \neq \delta(A|B)$; instead, a symmetric version of discord has been proposed by using measurements on both subsystems [167, 169]. The symmetric classical mutual information is

$$J(A : B)_\rho := \max_{\{A_a, B_b\}} J(\{A_a\} : \{B_b\}) \quad (4.5)$$

$$:= \max_{\{A_a, B_b\}} H(\{p_a^A\}) + H(\{p_b^B\}) - H(\{p_{ab}^{AB}\}), \quad (4.6)$$

where $p_{ab}^{AB} = \text{Tr}_{AB}[(A_a \otimes B_b)\rho_{AB}]$ and p_a^A, p_b^B are the marginals. The corresponding discord quantifier is then:

Definition 13. The **symmetric discord** of a bipartite state ρ_{AB} is

$$\delta(A : B)_\rho := I(A : B)_\rho - J(A : B)_\rho \quad (4.7)$$

One of the main properties of discord, which opens up a number of ways to interpret it, is the set of states for which it vanishes:

Definition 14. A bipartite state ρ_{AB} is called **classical-quantum (CQ)** iff $\delta(B|A)_\rho = 0$; equivalently, iff it has the form

$$\rho_{AB} = \sum_a \lambda_a |\psi_a\rangle\langle\psi_a| \otimes \rho_{B|a}, \quad (4.8)$$

where $\{|\psi_a\rangle\}$ is an orthonormal basis for $\mathcal{H}_A$.

Similarly, ρ_{AB} is called **classical-classical (CC)** iff $\delta(A:B)_\rho = 0$; equivalently, iff

$$\rho_{AB} = \sum_{a,b} \lambda_{ab} |\psi_a\rangle\langle\psi_a| \otimes |\phi_b\rangle\langle\phi_b|, \quad (4.9)$$

where $\{|\phi_b\rangle\}$ is also an orthonormal basis for $\mathcal{H}_B$.

From this, one can see that a zero-discord state is such that there exists some local basis with the property that a local measurement on A in this basis does not perturb the global state. When such a basis exists, it is the eigenbasis of the local state ρ_A (an analogous statement is true for symmetric discord).

It is also clear that a CQ or CC state is necessarily separable, but the converse is not true. A simple example of a state with discord but no entanglement is $\frac{1}{2}|0\rangle\langle 0| \otimes |0\rangle\langle 0| + \frac{1}{2}|+\rangle\langle +| \otimes |1\rangle\langle 1|$. Therefore one often says that discord is a more general, or weaker, type of correlation than entanglement, in that it exists for a larger set of states. In fact, almost all states (in a measure-theoretic sense) have non-zero discord [152]. More quantitative statements can also be made by showing that every entanglement measure E has an associated discord-like measure Q_E such that $Q_E \geq E$ [170].

The measures of classical correlations have a number of important properties, as outlined in [43]. Namely, classical correlations vanish for product states, cannot increase under local operations (therefore are unchanged under local unitaries, which are reversible), and equal $S(\rho_A) = S(\rho_B)$ for pure states. Monotonicity under local operations is also shared by the total correlations $I(A:B)$. However, we cannot conclude that the difference between them, the discord, is similarly monotonic – this is the subject of the discussion in Section 4.3.3. It is at least true that $\delta(B|A)$ is non-increasing under local operations on the unmeasured system B [44].

4.2 Creating discord from coherence

4.2.1 Measures of multipartite coherence and discord

We first introduce the terminology needed to extend the resource theory of speakable coherence to multipartite cases. Suppose the whole system is composed of n subsystems A_k of finite dimension d_k , and that each has a preferred incoherent basis $\{|i_k\rangle \mid i_k = 0, \dots, d_k - 1\}$. For speakable coherence, it is often natural to assume [5, 118, 138, 150, 171] that the combined system has a corresponding incoherent basis given by tensor products of the individual incoherent basis states, i.e. $|\mathbf{i}\rangle := |i_1\rangle|i_2\rangle \dots |i_n\rangle$, where $\mathbf{i} = (i_1, i_2, \dots, i_n)$. Then the full set of incoherent states is

$$\mathcal{I}_{A_1 \dots A_n} := \left\{ \sum_{\mathbf{i}} p_{\mathbf{i}} |\mathbf{i}\rangle \langle \mathbf{i}| \right\}, \quad (4.10)$$

with the usual condition on the $p_{\mathbf{i}}$ being normalised probabilities. Alternatively, we might treat the subsystems asymmetrically by considering coherence to be a resource only on some set of m subsystems A_k , denoting the remaining subsystems by B_k . The relevant free states are then of the form

$$\mathcal{I}_{B_1 \dots B_{n-m} | A_1 \dots A_m} := \left\{ \sum_{\mathbf{i}=(i_1, \dots, i_m)} p_{\mathbf{i}} |\mathbf{i}\rangle \langle \mathbf{i}|_{A_1 \dots A_m} \otimes \sigma_{B_1 \dots B_{n-m} | \mathbf{i}} \right\}. \quad (4.11)$$

We call these **$\mathbf{A}_1 \dots \mathbf{A}_m$ -incoherent** states; in the bipartite case treated in [118], they are called **incoherent-quantum (IQ)**. For notational simplicity, from now on we contract the sets of A_k and B_k to $\mathbf{A}$ and $\mathbf{B}$. We also only use the definition (4.11), as the set (4.10) is a special case where there are no B_k .

Given a contractive pseudo-distance D , we can generalise the class of measures (3.29) to the multipartite setting:

$$C_D^{\mathbf{B}|\mathbf{A}}(\rho_{\mathbf{AB}}) := \min_{\sigma_{\mathbf{AB}} \in \mathcal{I}_{\mathbf{B}|\mathbf{A}}} D(\rho_{\mathbf{AB}}, \sigma_{\mathbf{AB}}). \quad (4.12)$$

As before, $C_D^{\mathbf{B}|\mathbf{A}}$ vanishes iff a state is **$\mathbf{A}$ -incoherent**. These measures are now monotonic under the set of **$\mathbf{A}$ -incoherent** (or incoherent-quantum) operations, which are defined by mapping $\mathcal{I}_{\mathbf{B}|\mathbf{A}}$ to itself [118]. (We work for now with the BCP construction, leaving aside the distinction between different classes of incoherent operations as discussed in Chapter 3.) For the case of the relative entropy, the closest **$\mathbf{A}$ -incoherent** state is given by dephasing in the incoherent basis [118], so

$$C_R^{\mathbf{B}|\mathbf{A}}(\rho_{\mathbf{AB}}) = S(\rho_{\mathbf{AB}} \| \Phi_{\mathbf{A}}(\rho_{\mathbf{AB}})), \quad (4.13)$$

where $\Phi_{\mathbf{A}}(\rho_{\mathbf{AB}}) = \sum_{\mathbf{i}}(|\mathbf{i}\rangle\langle\mathbf{i}|_{\mathbf{A}} \otimes I_{\mathbf{B}})\rho_{\mathbf{AB}}(|\mathbf{i}\rangle\langle\mathbf{i}|_{\mathbf{A}} \otimes I_{\mathbf{B}})$.

A strong indication of the connection between coherence and discord is given by the comparing the forms of the CQ (or CC) states and IQ states. Evidently, a state is IQ iff it is CQ and such that the basis $\{|\psi_{\mathbf{i}}\rangle\}$ coincides with the incoherent basis. Analogously to the construction of multipartite $\mathbf{A}$ -incoherent states, we define the set of **A-classical** states to be

$$\mathcal{C}_{\mathbf{B}|\mathbf{A}} := \left\{ \sum_{\mathbf{i}} p_{\mathbf{i}} |\psi_{\mathbf{i}}\rangle\langle\psi_{\mathbf{i}}|_{\mathbf{A}} \otimes \sigma_{\mathbf{B}|\mathbf{i}} \right\}, \quad (4.14)$$

where $|\psi_{\mathbf{i}}\rangle = |\psi_{i_1}^1\rangle|\psi_{i_2}^2\rangle\cdots|\psi_{i_m}^m\rangle$ such that $\{|\psi_{i_k}^k\rangle \mid i_k = 0, \dots, d_k - 1\}$ is a basis for A_k . A set of distance-based measures of discord can then be defined [151, 172, 173]:

$$\Delta_D^{\mathbf{B}|\mathbf{A}}(\rho_{\mathbf{AB}}) := \min_{\sigma_{\mathbf{AB}} \in \mathcal{C}_{\mathbf{B}|\mathbf{A}}} D(\rho_{\mathbf{AB}}, \sigma_{\mathbf{AB}}), \quad (4.15)$$

vanishing iff $\rho_{\mathbf{AB}}$ is $\mathbf{A}$ -classical.

Using the partition of subsystems $(A_1, \dots, A_m, \mathbf{B})$ in which all B_k are grouped together, we write the associated set of separable states as $\mathcal{S}_{A_1:\dots:A_m:\mathbf{B}}$. We can then quantify the statement that “discord goes beyond entanglement” by observing that $\mathcal{C}_{\mathbf{B}|\mathbf{A}} \subset \mathcal{S}_{A_1:\dots:A_m:\mathbf{B}}$, so [151, 172]

$$\begin{aligned} E_D^{A_1:\dots:A_m:\mathbf{B}}(\rho_{\mathbf{AB}}) &:= \min_{\sigma_{\mathbf{AB}} \in \mathcal{S}_{A_1:\dots:A_m:\mathbf{B}}} D(\rho_{\mathbf{AB}}, \sigma_{\mathbf{AB}}) \\ &\leq \Delta_D^{\mathbf{B}|\mathbf{A}}(\rho_{\mathbf{AB}}). \end{aligned} \quad (4.16)$$

A direct connection with the above coherence measures can be made. Suppose we denote the $\mathbf{A}$ -incoherent states with respect to some product basis $b = \{|\psi_{\mathbf{i}}\rangle\}$ by $\mathcal{I}_{\mathbf{B}|\mathbf{A},b}$. Similarly, we write the corresponding coherence measures as $C_D^{\mathbf{B}|\mathbf{A},b}$ and dephasing channel as $\Phi_{\mathbf{A},b}$. Then it is clear that

$$\mathcal{C}_{\mathbf{B}|\mathbf{A}} = \bigcup_b \mathcal{I}_{\mathbf{B}|\mathbf{A},b}, \quad (4.17)$$

$$\Delta_D^{\mathbf{B}|\mathbf{A}}(\rho_{\mathbf{AB}}) = \min_b C_D^{\mathbf{B}|\mathbf{A},b}(\rho_{\mathbf{AB}}). \quad (4.18)$$

See Fig. 4.1 for a depiction of the relations between these different sets and measures.

From the relative entropy coherence measure, it follows that

$$\Delta_R^{\mathbf{B}|\mathbf{A}}(\rho_{\mathbf{AB}}) = \min_b S(\rho_{\mathbf{AB}} \| \Phi_{\mathbf{A},b}(\rho_{\mathbf{AB}})). \quad (4.19)$$

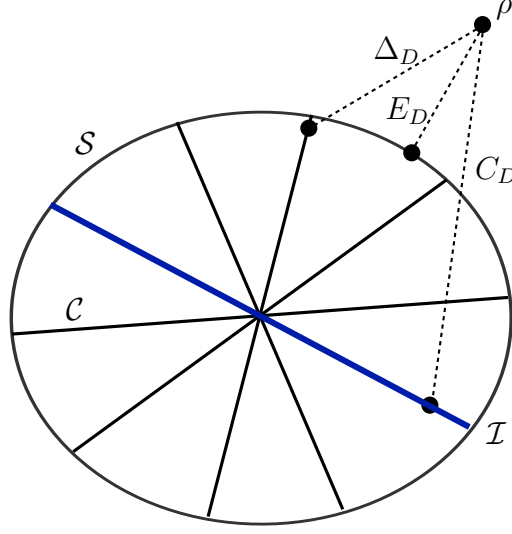


Figure 4.1: Schematic representation of the sets of separable states ($\mathcal{S}$, elliptical area), classical states ($\mathcal{C}$, solid lines) and incoherent states (line highlighted in blue), as well as the distance-based measures E_D , Δ_D , and C_D .

More in the spirit of the original discord quantity discussed in the previous section, one can define a generalisation of $\delta(B|A)$ to multipartite systems, known as **global discord** [174]. This relies on rewriting

$$\delta(B|A)_\rho = \min_b S(\rho_{AB} \| \Phi_{A,b}(\rho_{AB})) - S(\rho_A \| \Phi_{A,b}(\rho_A)), \quad (4.20)$$

which assumes that the minimisation over measurements on A is restricted to orthogonal projectors. This is extended to

$$\delta(\mathbf{B}|\mathbf{A})_\rho := \min_{\text{prod. } b} S(\rho_{\mathbf{AB}} \| \Phi_{\mathbf{A},b}(\rho_{\mathbf{AB}})) - \sum_{k=1}^m S(\rho_{A_k} \| \Phi_{A_k,b}(\rho_{A_k})), \quad (4.21)$$

where the minimisation is over all *product* basis choices. By a simple expansion and rearrangement of terms, this can be recast as

$$\delta(\mathbf{B}|\mathbf{A})_\rho = \min_{\text{prod. } b} I(A_1 : \dots : A_m : B_1 : \dots : B_{n-m})_\rho - I(A_1 : \dots : A_m : B_1 : \dots : B_{n-m})_{\Phi_{\mathbf{A},b}(\rho)}, \quad (4.22)$$

where $I(A : B : C : \dots) := S(\rho_A) + S(\rho_B) + S(\rho_C) + \dots - S(\rho_{ABC\dots}) = S(\rho_{ABC\dots} \| \rho_A \otimes \rho_B \otimes \rho_C \otimes \dots)$ is a multipartite version of the mutual information. The second term on the right hand side of (4.22) is a measure of classical correlations with respect to a measurement in the basis b on $\mathbf{A}$; thus, the whole expression is a generalisation of (4.4).

4.2.2 General bounds

In this section, we give general bounds for the conversion between coherence and discord. The setting is similar to that introduced in [138]. There, one has a system S and ancilla A , initially in a product state $\rho_S \otimes \rho_A$ and with ρ_A assumed to be incoherent. Their first observation is that, when ρ_S is also incoherent, a global incoherent operation always results in a separable state. When ρ_S contains coherence, a global incoherent operation can create entanglement between S and A . This relation can be made quantitative by giving an upper bound on the attainable entanglement from a certain amount of coherence. Moreover, it is shown there that this connection provides a valid coherence measure from any valid entanglement measure.

Here, we show that similar statements can be made for coherence and discord. We also generalise the setting to allow for multipartite systems as described above. We suppose that there are m subsystems A_k in which coherence is considered a resource, and $n - m$ remaining subsystems B_k . The initial state is assumed to be completely uncorrelated:

$$\rho_{\mathbf{AB}} = \bigotimes_{k=1}^m \rho_{A_k} \otimes \bigotimes_{k=1}^{n-m} \rho_{B_k}. \quad (4.23)$$

A global $\mathbf{A}$ -incoherent channel $\mathcal{E}$ is then applied to produce the final state $\mathcal{E}(\rho_{\mathbf{AB}})$ (see Fig. 4.2).

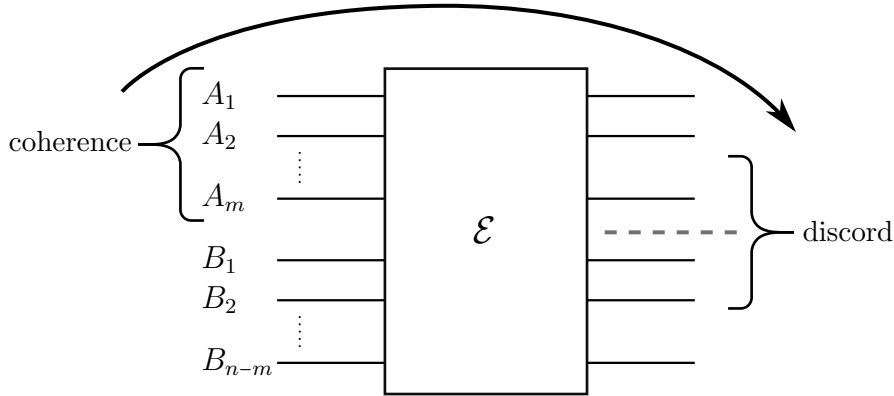


Figure 4.2: Converting coherence in subsystems $A_1, \dots, A_m$ into discord between $\mathbf{A}$ and $\mathbf{B}$ using a global $\mathbf{A}$ -incoherent channel $\mathcal{E}$. The initial state is fully uncorrelated.

Result 7. *For any contractive pseudo-distance D and initial product state (4.23), the discord created by an $\mathbf{A}$ -incoherent channel $\mathcal{E}$ is bounded by*

$$\Delta_D^{\mathbf{B}|\mathbf{A}}(\mathcal{E}(\rho_{\mathbf{AB}})) \leq C_D(\rho_{\mathbf{A}}). \quad (4.24)$$

(Recall the definition (3.29) of $C_D(\rho_A) = \min_{\sigma_A \in \mathcal{I}_A} D(\rho_A, \sigma_A)$.)

If D is the relative entropy or the Bures distance, and if the dimensions of $\mathbf{A}$, $\mathbf{B}$ satisfy $d_B \geq d_A$, then there is always a choice of ρ_B and $\mathcal{E}$ that saturates the inequality.

Proof. Suppose that σ_B is the closest incoherent state to ρ_A , such that $C_D(\rho_A) = D(\rho_A, \sigma_A)$. Two observations are useful here: firstly, note that

$$D(\rho_A \otimes \rho_B, \sigma_A \otimes \rho_B) = D(\rho_A, \sigma_A), \quad (4.25)$$

which holds because the operation of adding an uncorrelated system in the state ρ_B is reversible by tracing out $\mathbf{B}$ (so that monotonicity holds in both directions). Secondly, $\sigma_A \otimes \rho_B$ is $\mathbf{A}$ -incoherent, so by the assumption on $\mathcal{E}$,

$$\sigma_A \otimes \rho_B \in \mathcal{I}_{B|A} \Rightarrow \mathcal{E}(\sigma_A \otimes \rho_B) \in \mathcal{I}_{B|A} \subset \mathcal{C}_{B|A}. \quad (4.26)$$

Thus, we have the chain of relations

$$\begin{aligned} C_D(\rho_A) &= D(\rho_A, \sigma_A) \\ &= D(\rho_A \otimes \rho_B, \sigma_A \otimes \rho_B) \\ &\geq D(\mathcal{E}(\rho_A \otimes \rho_B), \mathcal{E}(\sigma_A \otimes \rho_B)) \\ &\geq \min_{\tau_{AB} \in \mathcal{C}_{B|A}} D(\mathcal{E}(\rho_A \otimes \rho_B), \tau_{AB}) \\ &= \Delta_D^{B|A}(\rho_{AB}). \end{aligned} \quad (4.27)$$

For the second statement, we assume for simplicity that $d_B = d_A$ (the case $d_B > d_A$ follows trivially). We choose an arbitrary basis $\{|i\rangle_B\}$ for $\mathbf{B}$, which is not necessarily a product basis but is labelled in the same way as the incoherent basis for $\mathbf{A}$. Then we choose the initial state of $\mathbf{B}$ to be $\rho_B = |0\rangle\langle 0|$. The $\mathbf{A}$ -incoherent operation is given by a unitary U which generalises the C-NOT to higher dimensions. Without writing out the gate in full, what is important is that it has the action

$$U|i\rangle_A|0\rangle_B = |i\rangle_A|i\rangle_B. \quad (4.28)$$

It was shown in [138] that in this case, the amount of *bipartite* entanglement generated is $E_D^{A:B}(U[\rho_A \otimes \rho_B]U^\dagger) = C_D(\rho_A)$, when D is the relative entropy or the Bures distance. Since every $\mathbf{A}$ -classical state is separable with respect to the $\mathbf{A} : \mathbf{B}$ partition,

i.e. $\mathcal{C}_{B|A} \subset \mathcal{S}_{A:B}$, we must have

$$\begin{aligned}\Delta_D^{B|A}(U[\rho_A \otimes \rho_B]U^\dagger) &\geq E_D^{A:B}(U[\rho_A \otimes \rho_B]U^\dagger) \\ &= C_D(\rho_A).\end{aligned}\tag{4.29}$$

It follows that the upper bound (4.24) must be saturated. (When $d_B > d_A$, there are additional basis states for B which do not enter into the calculation.) $\square$

Since discord is generally greater than entanglement, $\Delta_D^{B|A} \geq E_D^{A:B}$, our bound is tighter than the entanglement bound in [138]. Hence it is natural to ask whether coherence can be converted into discord without creating any entanglement. A simple two-qubit example is given by taking the initial state $|+\rangle\langle+|_A \otimes |0\rangle\langle 0|_B$ and the incoherent global operation

$$\mathcal{E}(\rho_{AB}) = pU\rho_{AB}U^\dagger + \left(\frac{1-p}{4}\right)I \otimes I,\tag{4.30}$$

where $p \in [0, 1]$ and U is the C-NOT gate. The resulting state is $p|\phi^+\rangle\langle\phi^+| + (1-p)/4I \otimes I$, known as a Werner state ($|\phi^+\rangle$ is a Bell state). For $p = 1$, this is maximally entangled, while for $p \leq 1/3$ it is separable. In contrast, there is non-zero discord for all $p > 0$ [42].

In this example, both entanglement and discord reach their maximal values at $p = 1$. However, it is unclear whether the upper bound for discord may be saturated even with a separable final state.

Now we give a bound for the global discord measure $\delta(B|A)$. For this quantity, the coherence of individual subsystems now plays an important role.

Result 8. *For any initial product state (4.23), the global discord created by an A -incoherent operation $\mathcal{E}$ is bounded by*

$$\delta(B|A)_{\rho'} \leq \sum_{k=1}^m C_R(\rho_{A_k}) - C_R(\rho'_{A_k}),\tag{4.31}$$

where $\rho'_{AB} = \mathcal{E}(\rho_{AB})$. This is saturated when the optimal basis in the definition of $\delta(B|A)$ coincides with the incoherent basis, and $\mathcal{E}$ is unitary.

Proof. The key observation is that the right-hand side of (4.21) involves a minimisa-

tion over all basis choices, and by restricting to the incoherent basis, we have

$$\begin{aligned}\delta(\mathbf{B}|\mathbf{A})_{\rho'} &\leq S(\rho'_{\mathbf{AB}}\|\Phi_{\mathbf{A}}(\rho'_{\mathbf{AB}})) - \sum_{k=1}^m S(\rho'_{A_k}\|\Phi_{A_k}(\rho'_{A_k})) \\ &= C_R^{\mathbf{B}|\mathbf{A}}(\rho'_{\mathbf{AB}}) - \sum_{k=1}^m C_R(\rho'_{A_k}).\end{aligned}\quad (4.32)$$

We then use the fact that $\mathcal{E}$ is $\mathbf{A}$ -incoherent, which implies $C_R^{\mathbf{B}|\mathbf{A}}(\rho'_{\mathbf{AB}}) \leq C_R^{\mathbf{B}|\mathbf{A}}(\rho_{\mathbf{AB}})$, so

$$\delta(\mathbf{B}|\mathbf{A})_{\rho'} \leq C_R^{\mathbf{B}|\mathbf{A}}(\rho_{\mathbf{AB}}) - \sum_{k=1}^m C_R(\rho'_{A_k}). \quad (4.33)$$

Finally, the initial product structure means that $C_R^{\mathbf{B}|\mathbf{A}}(\rho_{\mathbf{AB}}) = \sum_{k=1}^m C_R(\rho_{A_k})$, from which the result follows. The inequality is saturated exactly when there is equality in the first line of (4.32), i.e. when the incoherent basis is the optimal one for measurement, and when the global coherence $C_R^{\mathbf{B}|\mathbf{A}}$ is unchanged. $\square$

This shows that, in order to create global discord, the local coherence of subsystems must be consumed.

Some comments can be made about both above results. Firstly, if the discord quantities were defined by optimising over all local measurements (not just projective local measurements), then in general lower values would be achieved and the upper bounds would still apply. In addition, since the global incoherent operations assumed here are the largest possible set, the upper bounds are also valid for any other class of incoherent operations. However, the saturation of the bound in Result 7 assumes that the generalised C-NOT gate is a free operation.

The optimal value in the bound of Result 8 is $\sum_{k=1}^m C_R(\rho_{A_k}) = C_R(\rho_{\mathbf{A}})$, the same as the bound of Result 7. This requires that each final local state ρ_{A_k} is incoherent. It is difficult to say in general when this value is achieved; however, it is always possible in the bipartite case, which reduces to the discord $\delta(B|A)$. To see this, note from (4.2) and (4.3) that the maximal classical mutual information $J(B|A)$ is achieved by minimising $p_a \sum_a S(\rho_{B|a})$ over all possible measurements on A . By choosing $\rho_B = |0\rangle\langle 0|$ and applying the generalised C-NOT gate U as in Result 7 (again assuming $d_B \geq d_A$), we have the final state

$$\rho'_{AB} = \sum_{i,j} \rho_{ij} |i\rangle\langle j|_A \otimes |i\rangle\langle j|_B, \quad (4.34)$$

where $\rho_{ij} = \langle i | \rho_A | j \rangle$. Choosing to measure A in the incoherent basis, we have $\rho_{B|i} = |i\rangle\langle i|$, with zero entropy. Therefore this is the optimal basis. The reduced states are $\rho'_A = \sum_i \rho_{ii} |i\rangle\langle i|_A$, $\rho'_B = \sum_i \rho_{ii} |i\rangle\langle i|_B$. It is easily seen that $J(B|A) = S(\rho'_B)$ and hence $\delta(B|A) = S(\rho'_A) - S(\rho'_{AB}) = H(\{\rho_{ii}\}) - S(\rho_A) = C_R(\rho_A)$, as claimed.

4.2.3 Application to DQC1 algorithm

One of the first important studies of the usefulness of discord in quantum information processing is [45], which looks at its role in a model of quantum computation called **deterministic quantum computation with one quantum bit (DQC1)** [163]. It is known for computation with pure states that genuine multipartite entanglement between an unbounded number of qubits is necessary in order to have an exponential advantage over classical computation [34]. However, the situation is much less clear for mixed states. DQC1 provides a striking case of a highly efficient computation with very little quantumness.

The scheme is depicted in Fig. 4.3. Only one pure qubit $\rho_A = |0\rangle\langle 0|$ is permitted, while the remainder $B_1, \dots, B_n$ are initially maximally mixed, $\rho_B = I^{\otimes n}/2^n$. A Hadamard gate is first applied to A , followed by a controlled-unitary (C-U) gate applying U to $\mathbf{B}$ conditioned on the state of A : $|0\rangle\langle 0| \otimes I^{\otimes n} + |1\rangle\langle 1| \otimes U$. A subsequent measurement of A enables estimation of the normalised trace of U , $\langle \sigma_A^x + i\sigma_A^y \rangle = \text{Tr}(U)/2^n =: \tau$. The protocol is repeated a number of times to evaluate the expectation value to a desired precision. Hence τ is determined in bounded (i.e. $O(1)$) time, up to a constant factor depending on the precision.

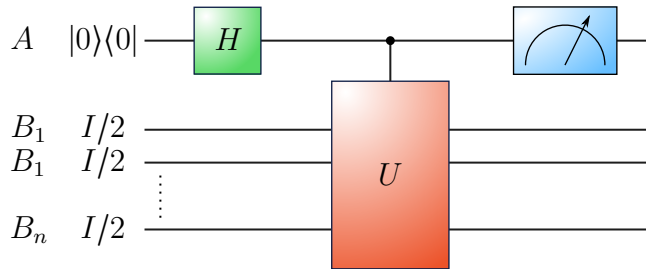


Figure 4.3: Circuit diagram for the DQC1 algorithm to evaluate the normalised trace $\tau = \text{Tr}(U)/2^n$ of a matrix U of dimension 2^n . The final measurement is not in the computational basis, but determines $\langle \sigma_A^x + i\sigma_A^y \rangle$.

Since U encodes a $2^n \times 2^n$ matrix (when written in some basis), it is plausible that its normalised trace τ may be hard to compute classically, meaning that it takes a

time exponential in n . Indeed, DQC1 accomplishes the computation exponentially faster than any known classical algorithm [163], and there is evidence that the model is hard to simulate on a classical computer [166, 175].

Therefore it is somewhat surprising that, with respect to the most natural bipartite splitting of A versus B , the state is separable at every stage of the algorithm. This is most easily seen by writing the unitary in a spectral decomposition as $U = \sum_{i=1}^{2^n} e^{i\phi_i} |i\rangle\langle i|$, so that the C-U gate can be rewritten as

$$|0\rangle\langle 0| \otimes I^{\otimes n} + |1\rangle\langle 1| \otimes U = \sum_{i=1}^{2^n} R_{\phi_i} \otimes |i\rangle\langle i|, \quad (4.35)$$

$$R_{\phi} := |0\rangle\langle 0| + e^{i\phi} |1\rangle\langle 1|. \quad (4.36)$$

The state before application of the C-U gate is $|+\rangle\langle +| \otimes \sum_{i=1}^{2^n} |i\rangle\langle i| / 2^n$, which then becomes

$$\rho'_{AB} = \sum_{i=1}^{2^n} \frac{1}{2^n} R_{\phi_i} |+\rangle\langle +| R_{\phi_i}^{\dagger} \otimes |i\rangle\langle i|, \quad (4.37)$$

which is explicitly separable. In addition, there is negligible entanglement in any other bipartite split (measured by the negativity) [176].

Looking for an alternative signature of quantumness, it was shown in [45] that the discord in the A - B split is generally non-zero. Hence it was conjectured that discord is the resource for the computation. We also observe that the C-U gate is A -incoherent, so coherence is only present as a result of the Hadamard gate acting on A . Result 8 can be directly applied to give $\delta(B|A)_{\rho'} \leq \Delta C_A := C_R(|+\rangle\langle +|) - C_R(\rho'_A)$. The final state of A is easily computed as

$$\rho'_A = \frac{1}{2} (|0\rangle\langle 0| + |1\rangle\langle 1| + \tau |1\rangle\langle 0| + \tau^* |0\rangle\langle 1|), \quad (4.38)$$

with eigenvalues $(1 \pm |\tau|)/2$. Therefore $S(\Phi(\rho'_A)) = 1$, $S(\rho'_A) = h([1 + |\tau|]/2)$, where $h(p) = -p \log p - (1-p) \log(1-p)$ is the binary Shannon entropy. Given that the state of A just after the Hadamard gate has $C_R = 1$, we have

$$\delta(B|A)_{\rho'} \leq \Delta C_A = h\left(\frac{1 + |\tau|}{2}\right). \quad (4.39)$$

Claims about discord being the signature of a quantum advantage in DQC1 have some problems. For example, for any $U = e^{i\phi} M$ such that $M^2 = I^{\otimes n}$, one finds that the discord vanishes [177]. This can be seen by examining (4.37). In this case we have $\phi_i = \phi$ or $\phi + \pi$, from which each term in the sum is of the form $R_{\phi} |\pm\rangle\langle \pm| R_{\phi}^{\dagger} \otimes |i\rangle\langle i|$, so

ρ'_{AB} is CQ. However, there is still no known efficient classical algorithm to determine τ for such a U .

On the other hand, coherence is generally still consumed when discord vanishes (the inequality is not generally saturated). We find that $\Delta C_A = 0$ iff $|\tau| = 1$, which is equivalent to $U = e^{i\phi} I^{\otimes n}$. Given the promise that U is of this form, it is trivial to determine τ classically – one just reads off the phase of a single diagonal element. This suggests that the consumption of coherence, rather than the creation of discord, might be the correct indicator of a quantum speed-up here.

This issue is, however, more subtle than it first appears. The difficulty is in how we compare the classical and quantum algorithms. The input for DQC1 is not in fact a matrix representation of U , but is typically considered to be a *classical description* of a sequence of elementary (one- and two-qubit) gates of polynomial length in n [176]. The polynomial requirement ensures that the circuit has an efficient implementation; the input must be classical (rather than a set of physical gates) so that we can compare the quantum algorithm with a classical one. So the claim we should like to make is: given a matrix representation of a polynomial sequence of elementary gates promised to result in an overall unitary $U = e^{i\phi} I^{\otimes n}$, there exists an efficient classical algorithm to determine ϕ . Unfortunately, it is not at all clear whether this is true, so it must be left as an open problem.

4.3 Incoherent operations and discord

4.3.1 Basis-dependent discord (BDD)

In this section, we change perspective somewhat in connecting coherence and discord. The main obstacle to understanding discord as a resource is the lack of a full characterisation of the operations under which it cannot increase. Unlike entanglement, discord can be created by local operations. For example, the classically correlated state $(|0\rangle\langle 0| \otimes |0\rangle\langle 0| + |1\rangle\langle 1| \otimes |1\rangle\langle 1|)/2$ can be converted into the discordant state $(|0\rangle\langle 0| \otimes |0\rangle\langle 0| + |+\rangle\langle +| \otimes |1\rangle\langle 1|)/2$ by a local channel on the first qubit with Kraus operators $K_0 = |0\rangle\langle 0|$, $K_1 = |+\rangle\langle 1|$. The full set of channels which cannot create discord from zero has been found [164, 178–180]. In that they are defined by preserving the set of free (classically correlated) states, these stand to discord as do the maximally incoherent channels (Definition 2) to coherence. Until these channels – or a subset

thereof – are given a physically sound motivation, we cannot have a resource theory of discord.

In order to make progress on this issue, we study a simpler version of discord where the measurement basis is fixed (and also focus on the bipartite case):

Definition 15. For a bipartite state ρ_{AB} and a fixed basis $b = \{|i\rangle_A | i = 0, \dots, d_A - 1\}$ for A , we define the **classical correlations with respect to b** as

$$J_b(B|A)_\rho := J(B|\{|i\rangle_A\langle i|_A\})_\rho = I(A : B)_{\Phi_{A,b}(\rho)}. \quad (4.40)$$

The corresponding **basis-dependent discord (BDD)** is

$$\delta_b(B|A)_\rho := I(A : B)_\rho - J_b(B|A)_\rho = I(A : B)_\rho - I(A : B)_{\Phi_{A,b}(\rho)}. \quad (4.41)$$

Analogously, when $b = \{|i\rangle_A |j\rangle_B | i = 0, \dots, d_A - 1; j = 0, \dots, d_B - 1\}$ specifies a product basis for AB , the symmetric versions of these quantities are

$$J_b(A : B) := J(\{|i\rangle_A\langle i|_A\} : \{|j\rangle_B\langle j|_B\}) = I(A : B)_{\Phi_{AB,b}(\rho)} \quad (4.42)$$

$$\delta_b(A : B) := I(A : B)_\rho - J_b(A : B)_\rho = I(A : B)_\rho - I(A : B)_{\Phi_{AB,b}(\rho)}. \quad (4.43)$$

For the remainder of this chapter, we take the basis b to be the incoherent basis. (Therefore we can drop the b subscript for the dephasing channel.) As we shall see, this lets us draw direct connections between BDD and coherence.

Operationally, BDD can be motivated as a quantity of interest by the following considerations: suppose, as in Chapter 3 when discussing the concept of “using coherence”, that the party in possession of A has access only to incoherent measurements. The accessible correlations between A and B are then $J_b(B|A)$. So $\delta_b(B|A)$ measures the excess in correlations contributing to the total $I(A : B)$. It is thus reasonable to describe BDD as measuring the quantumness of correlations with respect to the incoherent basis. This qualitative statement will be strengthened by the results presented below. Also note that the seminal paper [42] initially motivated the basis-dependent case before discussing $\delta(B|A) = \min_b \delta_b(B|A)$.

As an initial connection, we observe that BDD is the difference between measures of coherence in the global and local states:

$$\delta_b(B|A)_\rho = C_R^{B|A}(\rho_{AB}) - C_R(\rho_A), \quad (4.44)$$

$$\delta_b(A : B)_\rho = C_R^{AB}(\rho_{AB}) - C_R(\rho_A) - C_R(\rho_B), \quad (4.45)$$

where the proof is the same as for (4.20) and (4.21).

4.3.2 Vanishing BDD

We can gain some insight into the meaning of BDD by determining the set of states for which it vanishes. We know that the basis-independent discord $\delta(B|A)$ vanishes exactly for CQ states, which have the form $\sum_a \lambda_a |\psi_a\rangle\langle\psi_a| \otimes \rho_{B|a}$. Therefore it is tempting to guess that $\delta_b(B|A)$ vanishes only for IQ (or A -incoherent) states $\sum_i p_i |i\rangle\langle i| \otimes \rho_{B|i}$, since these are the CQ states restricted by setting $|\psi_i\rangle = |i\rangle$. In fact, we have the following:

Result 9. $\delta_b(B|A)_\rho = 0$ iff it is possible to write

$$\rho_{AB} = \sum_{\alpha} p_{\alpha} \rho_A^{\alpha} \otimes \rho_B^{\alpha} \quad \text{such that} \quad \text{Tr}[\Phi_A(\rho_A^{\alpha}) \Phi_A(\rho_A^{\alpha'})] = 0 \quad \forall \alpha \neq \alpha'. \quad (4.46)$$

Likewise, $\delta_b(A:B)_\rho = 0$ iff it is possible to write

$$\rho_{AB} = \sum_{\alpha, \beta} p_{\alpha\beta} |\psi_{\alpha}\rangle\langle\psi_{\alpha}| \otimes |\phi_{\beta}\rangle\langle\phi_{\beta}| \quad (4.47)$$

such that $\text{Tr}[\Phi_A(|\psi_{\alpha}\rangle\langle\psi_{\alpha}|) \Phi_A(|\psi_{\alpha'}\rangle\langle\psi_{\alpha'}|)] = 0 \quad \forall \alpha \neq \alpha'$ and sim. for $\Phi_B(|\phi_{\beta}\rangle\langle\phi_{\beta}|)$.

Before giving the proof, we recall an important theorem by Petz [181]:

Theorem 1 (Petz). *For two states ρ, σ , the relative entropy is unchanged under a quantum channel $\mathcal{E}$, meaning that $S(\rho||\sigma) = S(\mathcal{E}(\rho)||\mathcal{E}(\sigma))$, iff there exists a channel $\mathcal{R}$ satisfying $\mathcal{R} \circ \mathcal{E}(\rho) = \rho$, $\mathcal{R} \circ \mathcal{E}(\sigma) = \sigma$. Such an $\mathcal{R}$ is called a **recovery channel**. When recovery is possible, it is achieved by a channel with the explicit formula*

$$\mathcal{R}(X) = \sigma^{1/2} \mathcal{E}^{\dagger}[\mathcal{E}(\sigma)^{-1/2} X \mathcal{E}(\sigma)^{-1/2}] \sigma^{1/2}. \quad (4.48)$$

Proof of Result 9. We focus first on $\delta_b(B|A)$. For notational simplicity, we let $\sigma_{AB} = \Phi_A(\rho_{AB})$. Writing (4.41) in terms of the relative entropy, we see that

$$\delta_b(B|A)_\rho = 0 \Leftrightarrow S(\rho_{AB}||\rho_A \otimes \rho_B) = S(\sigma_{AB}||\sigma_A \otimes \rho_B). \quad (4.49)$$

Therefore we can apply Theorem 1 using the channel $\mathcal{E} = \Phi_A = \Phi_A^{\dagger}$. After writing

$$\sigma_{AB} = \sum_i p_i |i\rangle\langle i| \otimes \rho_{B|i}, \quad (4.50)$$

the recovery condition for ρ_{AB} , with $\mathcal{R}$ given by (4.48), says that

$$\begin{aligned}
\rho_{AB} &= \mathcal{R}(\sigma_{AB}) \\
&= \sum_i p_i \mathcal{R}(|i\rangle\langle i| \otimes \rho_{B,i}) \\
&= \sum_i p_i (\rho_A \otimes \rho_B)^{1/2} \Phi_A[(\sigma_A \otimes \rho_B)^{-1/2} (|i\rangle\langle i| \otimes \rho_{B,i}) (\sigma_A \otimes \rho_B)^{-1/2}] (\rho_A \otimes \rho_B)^{1/2} \\
&= \sum_i (\rho_A^{1/2} |i\rangle\langle i| \rho_A^{1/2}) \otimes \rho_{B,i}.
\end{aligned} \tag{4.51}$$

Since $\delta(B|A) \leq \delta_b(B|A) = 0$, ρ_{AB} must be CQ, so we can write

$$\rho_{AB} = \sum_{\alpha} \lambda_{\alpha} |\psi_{\alpha}\rangle\langle\psi_{\alpha}| \otimes \rho_{B|\alpha} = \sum_i (\rho_A^{1/2} |i\rangle\langle i| \rho_A^{1/2}) \otimes \rho_{B,i}, \tag{4.52}$$

where $\rho_A = \sum_{\alpha} \lambda_{\alpha} |\psi_{\alpha}\rangle\langle\psi_{\alpha}|$ is a spectral decomposition for ρ_A . Multiplying this equation by $\rho_A^{-1/2} \otimes I$ on the left and right gives

$$\sum_{\alpha} |\psi_{\alpha}\rangle\langle\psi_{\alpha}| \otimes \rho_{B|\alpha} = \sum_i |i\rangle\langle i| \otimes \rho_{B,i}. \tag{4.53}$$

Note that we have defined the inverse of ρ_A as acting only on its support:

$$\rho_A^{-1/2} := \sum_{\alpha: \lambda_{\alpha} \neq 0} \lambda_{\alpha}^{-1/2} |\psi_{\alpha}\rangle\langle\psi_{\alpha}|. \tag{4.54}$$

Therefore in (4.53), we have included only the terms that are in the support of ρ_A .

We now multiply (4.53) on the left by $\langle i|$ and on the right by $|\psi_{\alpha}\rangle$, giving $\langle i|\psi_{\alpha}\rangle \rho_{B|\alpha} = \langle i|\psi_{\alpha}\rangle \rho_{B,i} \forall \alpha, i$. Therefore either $\langle i|\psi_{\alpha}\rangle = 0$ or $\rho_{B|\alpha} = \rho_{B,i}$. This means that any two $|\psi_{\alpha}\rangle$ either correspond to the same $\rho_{B|\alpha}$ or are in the span of two disjoint sets of incoherent basis vectors – in which case the two dephased states $\Phi_A(|\psi_{\alpha}\rangle\langle\psi_{\alpha}|)$ are orthogonal. If we group together all the terms corresponding to the same $\rho_{B|\alpha}$, we arrive at the claimed form (4.46).

Conversely, we check via (4.44) that all states of this form do have vanishing $\delta_b(B|A)$. We have to show that $C_R^{B|A}(\rho_{AB}) = C_R(\rho_A)$. Firstly we calculate

$$\begin{aligned}
S(\rho_{AB}) &= S\left(\sum_{\alpha} p_{\alpha} \rho_A^{\alpha} \otimes \rho_B^{\alpha}\right) \\
&= H(\{p_{\alpha}\}) + \sum_{\alpha} p_{\alpha} [S(\rho_A^{\alpha}) + S(\rho_B^{\alpha})],
\end{aligned} \tag{4.55}$$

using the fact that different ρ_A^{α} are orthogonal. Similarly, writing $\sigma_A^{\alpha} := \Phi_A(\rho_A^{\alpha})$ and $\sigma_{AB} = \sum_{\alpha} p_{\alpha} \sigma_A^{\alpha} \otimes \rho_B^{\alpha}$,

$$S(\sigma_{AB}) = H(\{p_{\alpha}\}) + \sum_{\alpha} p_{\alpha} [S(\sigma_A^{\alpha}) + S(\rho_B^{\alpha})]. \tag{4.56}$$

It follows that

$$\begin{aligned}
C_R^{B|A}(\rho_{AB}) &= \left[H(\{p_\alpha\}) + \sum_\alpha p_\alpha S(\sigma_A^\alpha) \right] - \left[H(\{p_\alpha\}) + \sum_\alpha p_\alpha S(\rho_A^\alpha) \right] \\
&= S(\sigma_A) - S(\rho_A) \\
&= C_R(\rho_A).
\end{aligned} \tag{4.57}$$

For the symmetric discord, note that $J(A : B) \leq J(B|A)$ (since $\Phi_{AB}(\rho_{AB})$ is obtained from $\Phi_A(\rho_{AB})$ by an operation local to B), and so $\delta_b(A : B) \geq \delta_b(B|A)$. Hence $\delta_b(A : B) = 0 \Rightarrow \delta_b(B|A) = \delta_b(A|B) = 0$. By the above result, we can write

$$\rho_{AB} = \sum_\alpha p_\alpha |\psi_\alpha\rangle\langle\psi_\alpha| \otimes \rho_{B|\alpha} = \sum_\beta q_\beta \rho_{A|\beta} \otimes |\phi_\beta\rangle\langle\phi_\beta|, \tag{4.58}$$

where the $|\phi_\beta\rangle$ obey the same conditions as $|\psi_\alpha\rangle$. We multiply this equation by $\langle\psi_\alpha| \langle\phi_\beta|$ on the left and $|\psi_\alpha\rangle |\phi_{\beta'}\rangle$ on the right. If we take $\beta \neq \beta'$, this gives

$$p_\alpha \langle\phi_\beta| \rho_{B|\alpha} |\phi_{\beta'}\rangle = 0 \quad \forall \alpha, \beta \neq \beta'. \tag{4.59}$$

For all α such that $p_\alpha \neq 0$, this implies that $\rho_{B|\alpha}$ is diagonal in the basis $\{|\phi_\beta\rangle\}$, and can therefore be written as $\rho_{B|\alpha} = \sum_\beta p_{\beta|\alpha} |\phi_\beta\rangle\langle\phi_\beta|$. Therefore ρ_{AB} takes the form

$$\rho_{AB} = \sum_{\alpha, \beta} p_\alpha p_{\beta|\alpha} |\psi_\alpha\rangle\langle\psi_\alpha| \otimes |\phi_\beta\rangle\langle\phi_\beta|, \tag{4.60}$$

giving the claimed result by letting $p_{\alpha\beta} := p_\alpha p_{\beta|\alpha}$. $\square$

States with vanishing BDD can be coherent, but each coherent subspace in the expansion (4.46) is associated with a product state. Intuitively, in these states, coherence is stored independently of correlations. The condition on the different components ρ_A^α says that they are perfectly distinguishable by a measurement in the incoherent basis.

While every such state is A -incoherent, the converse is not true. A counterexample is given by the following qutrit-qubit state:

$$\rho_{AB} = \frac{1}{2} |+_01\rangle\langle+_01| \otimes |0\rangle\langle 0| + \frac{1}{2} |2\rangle\langle 2| \otimes |1\rangle\langle 1|, \tag{4.61}$$

where $|\pm_{ij}\rangle := (|i\rangle \pm |j\rangle)/\sqrt{2}$. Note that an incorrect claim of the converse is made in Equation (16) of the seminal paper [42]. We have been unable to find any correction of this in the literature.¹ A -incoherent states are characterised by vanishing $C_R^{B|A}$ [118],

while the remaining set of states with zero BDD have $C_R^{B|A}(\rho_{AB}) = C_R(\rho_A) > 0$. This also includes all product states. In the case of $d_A = 2$, every state with vanishing BDD is in fact either A -incoherent or a product state; more complex cases such as (4.61) can only exist for $d_A \geq 3$. See Fig. 4.4 for a representation of these various sets.

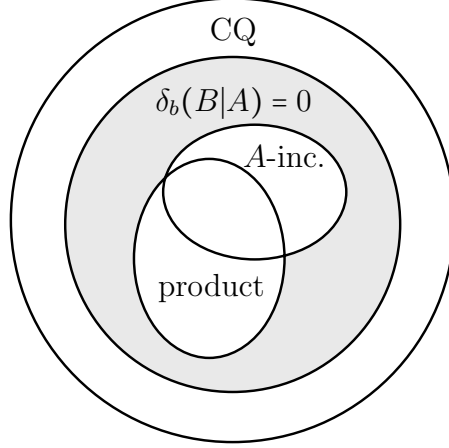


Figure 4.4: Representation of the sets of states for which the following quantities vanish: $\delta(B|A)$ (CQ states), $\delta_b(B|A)$ (Result 9), $C_R^{B|A}$ (A -incoherent states) and $I(A : B)$ (product states). This is schematic only, as not all these sets are convex. For $d_A = 2$, the shaded region is empty.

4.3.3 Behaviour under local operations

We now investigate the behaviour of correlations under local incoherent operations. We discuss both deterministic and ensemble monotonicity. The mutual information $I(A : B)$, as a measure of total correlations, is well known to be a deterministic and ensemble monotone under arbitrary local operations [165]. Therefore one may expect a similar result to hold for classical correlations under some local incoherent operations; we have the following:

Result 10. *$J_b(B|A)$ is a deterministic and ensemble monotone under local SI operations on A , as well as under arbitrary operations on B . $J_b(A : B)$ is similarly a monotone under local SI operations on both A and B .*

The set of SI operations is the largest subset of incoherent operations for which this statement is true.

¹One may ask where the mistake is in ‘Proof 3’ on page 4 of [42]. We suspect the error is in the step where an arbitrary off-diagonal and its complex conjugate are deleted from the density matrix. This does not in general result in a valid state (i.e. satisfying $\rho \geq 0$), and the authors have at least not shown that it does in this case.

Proof. First, let $\{\mathcal{E}_\mu\}$ be an SI quantum instrument, and let $\mathcal{F}(\rho) = \sum_\mu |\mu\rangle\langle\mu|_X \otimes \mathcal{E}_{\mu_A}(\rho)$ be the channel which performs the instrument with a classical memory X . Applying this to an initial state ρ_{AB} , we find

$$\begin{aligned} J(B|A)_\rho &= S(\Phi_A(\rho_{AB}) || \Phi_A(\rho_A) \otimes \rho_B) \\ &\geq S(\mathcal{F}_A \circ \Phi_A(\rho_{AB}) || \mathcal{F}_A \circ \Phi_A(\rho_A) \otimes \rho_B) \\ &= S(\Phi_A \circ \mathcal{F}_A(\rho_{AB}) || \Phi_A \circ \mathcal{F}_A(\rho_A) \otimes \rho_B) \\ &= S\left(\sum_\mu |\mu\rangle\langle\mu|_X \otimes \Phi_A \circ \mathcal{E}_{\mu_A}(\rho_{AB}) \left\| \sum_\mu |\mu\rangle\langle\mu|_X \otimes \mathcal{E}_{\mu_A}(\rho_A) \otimes \rho_B\right.\right), \end{aligned} \quad (4.62)$$

having used the contractivity of the relative entropy and the fact that $[\mathcal{F}_A, \Phi_A] = 0$. The relative entropy is also additive for direct sums [165], giving

$$J_b(B|A)_\rho \geq \sum_\mu p_\mu S(|\mu\rangle\langle\mu|_X \otimes \Phi_A(\rho_{\mu_{AB}}) || |\mu\rangle\langle\mu|_X \otimes \Phi_A(\rho_{\mu_A} \otimes \rho_B)), \quad (4.63)$$

where $p_\mu \rho_{\mu_{AB}} = \mathcal{E}_{\mu_A}(\rho_{AB})$. Since the relative entropy is unchanged under tracing out an uncorrelated ancilla,

$$\begin{aligned} J_b(B|A)_\rho &\geq \sum_\mu p_\mu S(\Phi_A(\rho_{\mu_{AB}}) || \Phi_A(\rho_{\mu_A}) \otimes \rho_B) \\ &= \sum_\mu p_\mu J(B|A)_{\rho_\mu}. \end{aligned} \quad (4.64)$$

The same calculation is easily seen to carry through for an arbitrary operation on B , or if we replace $J_b(B|A)$ by $J_b(A : B)$.

Conversely, let $\mathcal{E}$ be an incoherent operation such that $[\mathcal{E}, \Phi_A] \neq 0$. This means that there exist i, j, k with $i \neq j$ such that $\langle k | \mathcal{E}(|i\rangle\langle j|) | k \rangle \neq 0$. Let $\sigma_i := \mathcal{E}(|i\rangle\langle i|)$, $\sigma_j := \mathcal{E}(|j\rangle\langle j|)$, $\tau := \mathcal{E}(|i\rangle\langle j|)$. We construct an initial state with no classical correlations but which has $J_b(B|A) > 0$ after applying $\mathcal{E}$ to A . The state is

$$\rho_{AB} = \frac{1}{2} \left(|\phi\rangle\langle\phi| \otimes |0\rangle\langle 0| + \frac{1}{2} [|i\rangle\langle i| + |j\rangle\langle j|] \otimes |1\rangle\langle 1| \right), \quad (4.65)$$

where $|\phi\rangle := (|i\rangle + e^{i\phi}|j\rangle)/\sqrt{2}$. There are no classical correlations in this state, as

$$\Phi_A(\rho_{AB}) = \frac{1}{2} (|i\rangle\langle i| + |j\rangle\langle j|) \otimes \frac{1}{2} (|0\rangle\langle 0| + |1\rangle\langle 1|). \quad (4.66)$$

Applying $\mathcal{E}_A$, the state becomes

$$\rho'_{AB} := \mathcal{E}_A(\rho_{AB}) = \frac{1}{2} \left(\frac{1}{2} [\sigma_i + \sigma_j + e^{-i\phi}\tau + e^{i\phi}\tau^\dagger] \otimes |0\rangle\langle 0| + \frac{1}{2} [\sigma_i + \sigma_j] \otimes |1\rangle\langle 1| \right) \quad (4.67)$$

$$\sigma'_{AB} := \Phi_A(\rho'_{AB}) = \frac{1}{2} (\sigma_i + \sigma_j + \Phi_A[e^{-i\phi}\tau + e^{i\phi}\tau^\dagger]) \otimes |0\rangle\langle 0| + \frac{1}{2} (\sigma_i + \sigma_j) \otimes |1\rangle\langle 1|. \quad (4.68)$$

We see that $J(B|A)_{\rho'} = 0$ iff $\Phi_A(e^{-i\phi}\tau + e^{i\phi}\tau^\dagger) = 0$. However, by assumption, there exists k such that $\langle k|\tau|k\rangle \neq 0$. Then a value of ϕ can be chosen such that

$$\langle k|e^{-i\phi}\tau + e^{i\phi}\tau^\dagger|k\rangle = 2\cos\phi \operatorname{Re}[\langle k|\tau|k\rangle] + 2\sin\phi \operatorname{Im}[\langle k|\tau|k\rangle] \neq 0. \quad (4.69)$$

In this case, $J_b(B|A)_{\rho'} > J_b(B|A)_\rho = 0$, so classical correlations have been created. $\square$

Given the characterisation in Chapter 3 that SI operations do not use coherence, this result is not surprising: local SI operations cannot convert coherence into correlations in the incoherent basis.

We now look at BDD; as a measure of quantumness of correlations with respect to the incoherent basis, this might be expected to be a monotone under some suitably classical local operations. A useful trick [182] in the proof is to write $\delta_b(B|A)$ in terms of the **quantum conditional mutual information**, defined as

$$\begin{aligned} I(A : B|C)_\rho &:= I(AC : B)_\rho - I(C : B)_\rho \\ &= I(A : BC)_\rho - I(A : C)_\rho. \end{aligned} \quad (4.70)$$

Let C be a system which makes a measurement of A in the incoherent basis, such that after tracing out A , the state is $\sum_i \rho_{ii}|i\rangle\langle i|_C \otimes \rho_{B|i}$, where ρ_{ii} are the diagonals of ρ_A . This can be accomplished by a unitary U between A and C such as the generalised C-NOT discussed in Result 7. If we let $\rho'_{ABC} := U_{AC}(\rho_{AB} \otimes |0\rangle\langle 0|_C)U_{AC}^\dagger$, then

$$\begin{aligned} \delta_b(B|A)_\rho &= I(A : B)_\rho - I(A : B)_{\Phi_A(\rho)} \\ &= I(AC : B)_{\rho'} - I(C : B)_{\rho'} \\ &= I(A : B|C)_{\rho'}. \end{aligned} \quad (4.71)$$

Result 11. $\delta_b(B|A)$ is:

- (i) a deterministic and ensemble monotone under arbitrary local operations on B ;
- (ii) an ensemble monotone under local SI operations on A with subselection on individual Kraus operators;
- (iii) a deterministic and ensemble monotone under local GI operations on A ;

Proof. (i) The proof of the first part is the same as that in [182] for the minimised discord $\delta(B|A)$, using the quantum conditional mutual information. Note that (4.71) can be re-expressed as

$$\delta_b(B|A)_\rho = I(A : BC)_{\rho'} - I(A : C)_{\rho'}. \quad (4.72)$$

The result follows from the monotonicity of the first term $I(A : BC)_{\rho'}$ under local operations on B (which commute with the unitary U interacting A and C).

(ii) Let $\{\mathcal{E}_\mu\}$ be a quantum instrument such that $\mathcal{E}_\mu(\rho) = K_\mu \rho K_\mu^\dagger$ where $\{K_\mu\}$ is a set of SI Kraus operators. Let $\mathcal{F}$ be the channel performs this instrument while keeping a classical memory X of the outcome. We define

$$\sigma_{AB} := \Phi_A(\rho_{AB}), \quad (4.73)$$

$$\rho'_{XAB} := \mathcal{F}_A(\rho_{AB}) = \sum_\mu p_\mu |\mu\rangle\langle\mu|_X \otimes \mathcal{E}_{\mu_A}(\rho_{AB}) = \sum_\mu p_\mu |\mu\rangle\langle\mu|_X \otimes \rho_{\mu AB}, \quad (4.74)$$

$$\sigma'_{XAB} := \Phi_A(\rho'_{XAB}) = \mathcal{F}_A(\sigma_{AB}), \quad (4.75)$$

where the last equality follows from $[\mathcal{F}_A, \Phi_A] = 0$.

The final average BDD is

$$\sum_\mu p_\mu \delta_b(B|A)_{\rho_\mu} = \sum_\mu p_\mu [I(A : B)_{\rho_\mu} - I(A : B)_{\sigma_\mu}]. \quad (4.76)$$

It is easily verified that $I(XA : B)_{\rho'} = H(\{p_\mu\}) + \sum_\mu p_\mu I(A : B)_{\rho_\mu}$ and $I(XA : B)_{\sigma'} = H(\{p_\mu\}) + \sum_\mu p_\mu I(A : B)_{\sigma_\mu}$, so that

$$\sum_\mu p_\mu \delta_b(B|A)_{\rho_\mu} = I(XA : B)_{\rho'} - I(XA : B)_{\sigma'}. \quad (4.77)$$

By monotonicity of the mutual information under local operations, $I(XA : B)_{\rho'} \leq I(A : B)_\rho$. On the other hand, classical correlations are not lost under the map $\mathcal{F}$. This is because σ_{AB} can be recovered from σ'_{XAB} by a local operation. To show this, we write $\sigma_{AB} = \sum_i \rho_{ii} |i\rangle\langle i|_A \otimes \rho_{B,i}$ and $K_\mu = \sum_i c_i^\mu |f_\mu(i)\rangle\langle i|$, so that

$$\sigma'_{XAB} = \sum_{\mu,i} \rho_{ii} |c_i^\mu|^2 |\mu\rangle\langle\mu|_X \otimes |f_\mu(i)\rangle\langle f_\mu(i)|_A \otimes \rho_{B,i}. \quad (4.78)$$

We can find a recovery operation $\mathcal{T}$ mapping $XA \rightarrow A$ such that

$$\mathcal{T}: |\mu\rangle_X |f_\mu(i)\rangle_A \rightarrow |i\rangle_A \quad \forall i, \mu. \quad (4.79)$$

This is possible because of the assumption that K_μ is SI – so f_μ is invertible and hence i can be determined from knowledge of μ and $f_\mu(i)$. It follows that

$$\mathcal{T}_{XA}(\sigma'_{XAB}) = \sum_{\mu,i} \rho_{ii} |c_i^\mu|^2 |i\rangle\langle i|_A \otimes \rho_{B,i} = \sigma_{AB}, \quad (4.80)$$

as $\sum_\mu |c_i^\mu|^2 = 1 \ \forall i$ from $\sum_\mu K_\mu^\dagger K_\mu = I$.

Therefore there is a local channel each way mapping $\sigma_{AB} \leftrightarrow \sigma'_{XAB}$, and so $I(A : B)_\sigma = I(AX : B)_{\sigma'}$. From (4.77), we conclude that $\sum_\mu p_\mu \delta_b(B|A)_{\rho_\mu} \leq I(A : B)_\rho - I(A : B)_\sigma = \delta_b(B|A)_\rho$.

(iii) Ensemble monotonicity under local GI operations follows from (ii), since every GI operation is SI. A deterministic GI operation always preserves the incoherent part of the state, so $\Phi_A(\rho'_{AB}) = \Phi_A(\rho_{AB})$. Thus $J_b(B|A)$ is left unchanged, while $I(A : B)$ cannot increase. $\square$

The proof of SI ensemble monotonicity is closely related to the dilation picture, Result 3. Recall that this says that every SI operation can be constructed from an incoherent measurement by an ancilla, and incoherent unitaries conditioned on measurements performed on the ancilla. The incoherent unitaries cause a “scrambling” (permutation) of the incoherent basis. The main idea is that, as long as we retain a (classical) memory X of which measurement outcome on the ancilla was obtained, it is possible to recover all classical correlations by reversing the effect of the incoherent unitaries. This implies that $I(XA : B)_{\rho'} \leq I(A : B)_\rho$, $J_b(B|XA)_{\rho'} = J_b(B|A)_\rho$. See Fig. 4.5 for an illustration.

However, importantly, $\delta_b(B|A)$ is not a deterministic monotone under all SI operations. When the memory X is lost, different permutations may be mixed, resulting in a loss of classical correlations. For example, consider the state (4.61), which has zero BDD. Suppose we apply to subsystem A the SI channel with Kraus operators $K_0 = I/\sqrt{2}$, $K_1 = (|0\rangle\langle 0| + |1\rangle\langle 2| + |2\rangle\langle 1|)/\sqrt{2}$. This results in

$$\rho'_{AB} = \frac{1}{4} [(|+_{01}\rangle\langle +_{01}| + |+_{02}\rangle\langle +_{02}|) \otimes |0\rangle\langle 0| + (|1\rangle\langle 1| + |2\rangle\langle 2|) \otimes |1\rangle\langle 1|], \quad (4.81)$$

which has $\delta_b(B|A) = \frac{3}{4} \log 3 \approx 0.189 > 0$. With GI operations, no permutations are used and so no classical information is lost.

We may take the view that local SI operations can convert quantumness (as measured by coherence) into quantumness of correlations. If the initial state is either

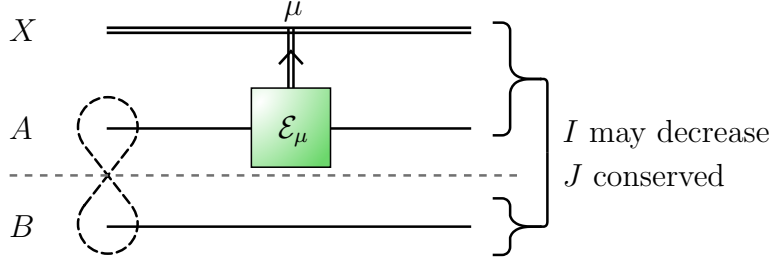


Figure 4.5: Circuit diagram showing a local SI operation $\mathcal{E} = \sum_{\mu} \mathcal{E}_{\mu}$ on a bipartite system AB . Each outcome μ corresponds to a sub-selected operation $\mathcal{E}_{\mu}(\rho) = K_{\mu}\rho K_{\mu}^{\dagger}$ with a single SI Kraus operator K_{μ} . The outcome is stored in a classical memory X local to A . Without discarding this memory, the total correlations $I(XA : B)$ cannot increase and the classical correlations $J_b(B|XA)$ are conserved.

A -incoherent or a product state, then clearly no local incoherent operation can create non-zero BDD. Therefore local SI creation of BDD from zero requires some initial coherence and correlations – this is impossible when A is a qubit. From (4.44), we can derive a quantitative bound under any local incoherent operation (just as in Section 4.2):

$$\delta_b(B|A)_{\rho'} - \delta_b(B|A)_{\rho} \leq C_R(\rho_A) - C_R(\rho'_A), \quad (4.82)$$

which shows that an increase in BDD comes at the expense of local coherence.

All of the above statements can be seen to hold analogously for the symmetric BDD $\delta_b(A : B)$, of course apart from part (i) of Result 11.

Finally, suppose that we permit classical communication between the two parties, meaning that each has access to the measurement outcomes obtained by the other. It should be clear that ensemble monotonicity as in part (ii) above still holds, where the sub-selection is based on the measurement outcomes on both sides.

4.3.4 Recovery from local dephasing

Apart from the characterisation given in Section 4.3.2, it is possible to expand upon the condition for BDD to vanish. From Theorem 1 and the proof of Result 9, we see that a state ρ_{AB} satisfies $\delta_b(B|A)_{\rho} = 0$ iff there exists a local operation $\mathcal{R}_A$ such that $\mathcal{R}_A \circ \Phi_A(\rho_{AB}) = \rho_{AB}$. In other words, the global state can be recovered by a local operation, after having lost all coherence locally. This is achieved by measuring subsystem A via projection onto the supports of $\Phi_A(\rho_A^{\alpha})$, and conditionally preparing ρ_A^{α} . Similarly, $\delta_b(A : B) = 0$ iff there exist $\mathcal{R}_A, \mathcal{R}_B$ such that

$(\mathcal{R}_A \circ \Phi_A) \otimes (\mathcal{R}_B \circ \Phi_B)(\rho_{AB}) = \rho_{AB}$. The parties at A and B simply project on the supports of $\Phi_A(|\psi_\alpha\rangle\langle\psi_\alpha|)$ and $\Phi_B(|\phi_\beta\rangle\langle\phi_\beta|)$ respectively, and prepare $|\psi_\alpha\rangle\langle\psi_\alpha|$ and $|\phi_\beta\rangle\langle\phi_\beta|$.

This gives a way of understanding that BDD is coherence shared between subsystems – and as such cannot be restored by a local operation. For non-zero BDD, it is natural to ask how to quantify the degree of recoverability. This could be expected to share some properties with $\delta_b(B|A)$.

Definition 16. *Let D be a contractive pseudo-distance. The **dephasing recoverability** (with asymmetric and symmetric versions) is defined as*

$$R_b^D(B|A)_\rho := \min_{\mathcal{R}_A} D(\rho_{AB}, \mathcal{R}_A \circ \Phi_A(\rho_{AB})), \quad (4.83)$$

$$R_b^D(A:B)_\rho := \min_{\mathcal{R}_A, \mathcal{R}_B} D(\rho_{AB}, (\mathcal{R}_A \circ \Phi_A) \otimes (\mathcal{R}_B \circ \Phi_B)(\rho_{AB})), \quad (4.84)$$

where the minimisation is performed over the set of local quantum channels.²

Note that there is a correspondence with the approach of [183] to (basis-independent) discord, in which recovery is attempted after an arbitrary local measurement. From the above discussion, $R_b^D(B|A)$ and $\delta_b(B|A)$ vanish concurrently, as do $R_b^D(A:B)$ and $\delta_b(A:B)$. In addition, we have the following monotonicity properties:

Result 12. $R_b^D(B|A)$ is:

- (i) a monotone under arbitrary local channels on B ;
- (ii) a monotone under local SI operations on A , retaining a classical memory X which records the measurement outcome obtained on the ancilla in the SI dilation;
- (iii) a deterministic monotone under local GI operations on A .

Proof. (i) Denote the channel on B as $\mathcal{E}_B$, and let the optimal recovery channel for the initial state be $\mathcal{R}_A^*$, such that $R(B|A)_\rho = D(\rho_{AB}, \mathcal{R}_A^*(\sigma_{AB}))$. (We use the same notation as in the proof of Result 11, with $\sigma_{AB} = \Phi_A(\rho_{AB})$.) With the final state

²Infimum and minimum are equivalent here, as the set of quantum channels is compact and so the infimum is always attainable.

$\rho'_{AB} = \mathcal{E}_B(\rho_{AB})$, we have

$$\begin{aligned}
R_b^D(B|A)_{\rho'} &\leq D(\rho'_{AB}, \mathcal{R}_A^*(\sigma'_{AB})) \\
&= D(\mathcal{E}_B(\rho_{AB}), \mathcal{R}_A^* \circ \mathcal{E}_B(\sigma_{AB})) \\
&= D(\mathcal{E}_B(\rho_{AB}), \mathcal{E}_B \circ \mathcal{R}_A^*(\sigma_{AB})) \\
&\leq D(\rho_{AB}, \mathcal{R}_A^*(\sigma_{AB})) \\
&= R_b^D(B|A)_{\rho},
\end{aligned} \tag{4.85}$$

where the second inequality follows from the contractivity of D .

(ii) As in the proof of Result 11, we denote by $\mathcal{F}$ the channel which performs the SI operation on A while keeping a classical memory X of the ancilla measurement outcome. The memory is assumed to be possessed by the party at A , and so the local recovery is allow to act on X and A together. We also make use of the channel $\mathcal{T}$ from Result 11, which locally recovers the incoherent part of the state using the memory: $\mathcal{T}(\sigma'_{XAB}) = \sigma_{AB}$.

The remaining calculation follows part (i), again using the channel $\mathcal{R}^*$ which optimally locally recovers the initial state ρ_{AB} from dephasing. In addition, we choose the particular recovery channel $\mathcal{R} := \mathcal{F} \circ \mathcal{R}^* \circ \mathcal{T}$ for the final state. This gives

$$\begin{aligned}
R_b^D(B|AX)_{\rho'} &\leq D(\mathcal{F}(\rho_{AB}), \mathcal{R}(\sigma'_{AB})) \\
&= D(\mathcal{F}(\rho_{AB}), \mathcal{F} \circ \mathcal{R}^* \circ \mathcal{T}(\sigma'_{AB})) \\
&= D(\mathcal{F}(\rho_{AB}), \mathcal{F} \circ \mathcal{R}^*(\sigma_{AB})) \\
&\leq D(\rho_{AB}, \mathcal{R}^*(\sigma_{AB})) \\
&= R_b^D(B|A)_{\rho}.
\end{aligned} \tag{4.86}$$

(iii) This case is simpler, as a local GI channel is such that $\sigma'_{AB} = \sigma_{AB}$. Then $R_b^D(B|A)_{\rho'} \leq D(\mathcal{E}_A(\rho_{AB}), \mathcal{E}_A \circ \mathcal{R}_A^*(\sigma'_{AB})) \leq D(\rho_{AB}, \mathcal{R}_A^*(\sigma_{AB})) = R_b^D(B|A)_{\rho}$. $\square$

As in the previous section, parts (ii) and (iii) also hold analogously for $R_b^D(A : B)$ under operations on B . In addition, the same comments regarding classical communication still apply. Specifically, let the two parties record their measurement outcomes

in the local classical memories X and Y respectively. Classical communication then simply entails copying information between X and Y . The crucial step in the proof of part (ii) – that there exists a local channel $\mathcal{T}$ recovering the σ_{AB} from σ'_{XYAB} – is still valid because no classical information has been lost.

4.3.5 Implications for basis-independent discord

In this section, we attempt to link coherence and standard basis-independent discord via their respective free operations. While there is currently no resource theory of discord, the full set of channels which cannot create discord is known to be the set of **commutativity-preserving** channels [178–180]. These are defined to be such that commuting pairs of states always remain commuting, i.e. $\mathcal{E}$ is commutativity preserving iff $[\mathcal{E}(\rho), \mathcal{E}(\sigma)] = 0$ for any pair ρ, σ such that $[\rho, \sigma] = 0$.

The commutativity-preserving operations can be divided into two classes. Those in the first class are known as **semiclassical** [164], and comprise all channels which completely dephase in some basis. A local semiclassical channel always outputs a classical-quantum state, thus destroying discord. Hence such channels are of limited interest.

The second class is the following [178–180]:

Definition 17. *In d dimensions, a channel $\mathcal{E}$ is called **isotropic** iff it is of the form*

$$\mathcal{E}(\rho) = pU\rho U^\dagger + (1-p)\frac{I}{d} \quad (\text{unitary-isotropic}) \quad (4.87)$$

or

$$\mathcal{E}(\rho) = pU\rho^T U^\dagger + (1-p)\frac{I}{d} \quad (\text{antiunitary-isotropic}), \quad (4.88)$$

where U is unitary and $p \in \mathbb{R}$ is such that $\mathcal{E}$ is CP.

For $d = 2$, this set consists of all unital channels [72], i.e. those preserving the identity, $\mathcal{E}(I) = I$. Generally, it is tempting to interpret the unitary-isotropic channels as being constructed from unitaries and mixing with the identity; this is not quite correct, as p may take on negative values (see Lemma 2 below) and so is not always a probability. On the other hand, while the operation $\rho \rightarrow \rho^T$ is not a CP map [72], it is possible to obtain a CP map with enough additional noise. However, the range of allowed p values for the antiunitary-isotropic channels becomes vanishing for

increasing d [180]. This may give an argument to doubt the physical relevance of such channels as free operations.

Recall that a state is free for discord (classical-quantum) iff there exists a product basis in which it is incoherent. On the other hand, given the rather restricted nature of the commutativity-preserving operations, it is natural to suppose a connection with operations that are incoherent with respect to *every* basis. In fact, we have the following:

Result 13. (i) *The only channel which is GI in every basis is the identity.*

(ii) *A channel $\mathcal{E}$ is SI in every basis iff it is **depolarising**, i.e. of the form $\mathcal{E}(\rho) = p\rho + (1-p)I/d$.*

As a corollary, a channel $\mathcal{E}$ is unitary-isotropic iff $\mathcal{E} = \mathcal{U} \circ \mathcal{F}$, where $\mathcal{F}$ is SI in every basis and $\mathcal{U}$ is unitary.

For the proof, we first establish some notation. We denote the set of incoherent states in basis b by $\mathcal{I}_b$. Similarly, we use $\mathcal{SI}_b$, $\mathcal{GI}_b$ and $\mathcal{MI}_b$ for the sets of SI, GI and maximal incoherent operations with respect to b , respectively. $\mathcal{D}$ is the set of depolarising channels.

Let us first prove part (i). The definition of GI operations can be written as

$$\mathcal{E} \in \mathcal{GI}_b \Leftrightarrow \forall \rho \in \mathcal{I}_b, \mathcal{E}(\rho) = \rho. \quad (4.89)$$

For an arbitrary state ρ , if we choose b to be its eigenbasis, then trivially $\rho \in \mathcal{I}_b$. Hence every $\mathcal{E} \in \mathcal{GI}_b$ must satisfy $\mathcal{E}(\rho) = \rho$.

Some Lemmas will be useful in establishing statement (ii).

Lemma 2. *In dimension d , the set of allowed values for p in the set $\mathcal{D}$ is*

$$\frac{-1}{d^2 - 1} \leq p \leq 1. \quad (4.90)$$

Proof. We first introduce the entangled states on two copies of the system $|\alpha_{kl}\rangle := \sum_{j=0}^{d-1} \frac{1}{\sqrt{d}} \omega^{jk} |j\rangle |j \oplus l\rangle$, where $k, l \in \{0, 1, \dots, d-1\}$, $\omega := e^{2\pi i/d}$ and $j \oplus l := j + l \mod d$. One can check that these form an orthonormal basis for the total dimension- d^2 Hilbert space.

Recall that a map $\mathcal{E}$ is CP iff $0 \leq \sigma := (\mathcal{E} \otimes I)(|\alpha_{00}\rangle\langle\alpha_{00}|)$ [72]. It is straightforward to see that for $\mathcal{E} \in \mathcal{D}$,

$$\begin{aligned}\sigma &= p|\alpha_{00}\rangle\langle\alpha_{00}| + \frac{(1-p)}{d^2}I \otimes I \\ &= p|\alpha_{00}\rangle\langle\alpha_{00}| + \frac{(1-p)}{d^2} \sum_{k,l} |\alpha_{kl}\rangle\langle\alpha_{kl}| \\ &= \left(\frac{1 + (d^2 - 1)p}{d^2} \right) |\alpha_{00}\rangle\langle\alpha_{00}| + \frac{(1-p)}{d^2} \sum_{(k,l) \neq (0,0)} |\alpha_{kl}\rangle\langle\alpha_{kl}|.\end{aligned}\quad (4.91)$$

This is a spectral decomposition for σ , and all eigenvalues are non-negative if and only if $p \in [-1/(d^2 - 1), 1]$. $\square$

Lemma 3. $\cap_b \mathcal{MI}_b$ is a subset of the unital channels.

Proof. The maximally mixed state I/d is the unique state which is incoherent in every basis. So for $\mathcal{E} \in \cap_b \mathcal{MI}_b$, $\mathcal{E}(I/d)$ must be incoherent in every basis, and thus equals I/d . $\square$

We proceed to use Lemmas 2 and 3 to prove the main statement, that $\cap_b \mathcal{SI}_b = \mathcal{D}$. We do this by proving the inclusions $\mathcal{D} \subseteq \cap_b \mathcal{SI}_b$ and $\cap_b \mathcal{SI}_b \subseteq \mathcal{D}$.

Firstly, to show that $\mathcal{D} \subset \mathcal{SI}_b \ \forall b$, we construct a set of SI Kraus operators in an arbitrary basis $b = \{|i\rangle\}_{i=0,\dots,d-1}$ for every operation in $\mathcal{D}$. For $k, l \in \{0, 1, \dots, d-1\}$, define

$$K_{kl} := \kappa_{kl} L_{kl}, \quad (4.92)$$

$$L_{kl} := \sum_{j=0}^{d-1} \omega^{jk} |j \oplus l\rangle\langle j|, \quad (4.93)$$

$$\kappa_{kl} := \begin{cases} \frac{\sqrt{1+(d^2-1)p}}{d} & k = l = 0, \\ \frac{\sqrt{1-p}}{d} & \text{otherwise.} \end{cases} \quad (4.94)$$

The operators K_{kl} are immediately seen to be SI in the chosen basis. By Lemma 2, the coefficients κ_{kl} are real since $p \in [-1/(d^2 - 1), 1]$. We find that $K_{kl}^\dagger K_{kl} = \kappa_{kl}^2 I$ and thus

$$\sum_{k,l} K_{kl}^\dagger K_{kl} = \left[\frac{1 + (d^2 - 1)p}{d^2} + \frac{(1-p)}{d^2} (d^2 - 1) \right] I = I, \quad (4.95)$$

as required for a trace-preserving map. Next we verify the action of the Kraus operators on the matrix elements of any state:

$$\begin{aligned}
\sum_{k,l} K_{kl} |j\rangle\langle j'| K_{kl}^\dagger &= K_{00} |j\rangle\langle j'| K_{00}^\dagger + \sum_{(k,l) \neq (0,0)} K_{kl} |j\rangle\langle j'| K_{kl}^\dagger \\
&= \frac{1 + (d^2 - 1)p}{d^2} L_{00} |j\rangle\langle j'| L_{00}^\dagger + \frac{(1-p)}{d^2} \left(\sum_{k,l} L_{kl} |j\rangle\langle j'| L_{kl}^\dagger - L_{00} |j\rangle\langle j'| L_{00}^\dagger \right) \\
&= \frac{1 + (d^2 - 1)p - (1-p)}{d^2} |j\rangle\langle j'| + \frac{(1-p)}{d^2} \sum_{k,l} \omega^{k(j-j')} |j \oplus l\rangle\langle j' \oplus l| \\
&= p |j\rangle\langle j'| + \frac{(1-p)}{d} \delta_{j,j'} \sum_l |j \oplus l\rangle\langle j' \oplus l| \\
&= p |j\rangle\langle j'| + \delta_{j,j'} (1-p) \frac{I}{d},
\end{aligned} \tag{4.96}$$

such that $\sum_{k,l} K_{kl} \rho K_{kl}^\dagger = p\rho + (1-p)I/d$ for any state ρ . This establishes that every element of $\mathcal{D}$ admits a set of SI Kraus operators in any basis.

Next, we take an arbitrary member $\mathcal{E}$ of $\cap_b \mathcal{SI}_b$ and show that it has the form of a depolarising channel. This proceeds in three stages. We show that the action on any pure state $|\psi\rangle$ is $\mathcal{E}(|\psi\rangle\langle\psi|) = p|\psi\rangle\langle\psi| + (1-p)I/d$, where p can be a function of $|\psi\rangle$. We then narrow this down to p being the same for all elements of a given basis, so that $\mathcal{E}(|i\rangle\langle i|) = p|i\rangle\langle i| + (1-p)I/d$ independent of i . Finally, we also deduce that $\mathcal{E}(|i\rangle\langle j|) = p|i\rangle\langle j|$ for any $i \neq j$. This suffices to determine the form of $\mathcal{E}$.

Take an arbitrary pure state $|\psi\rangle$, then there exists a basis $b = \{|i\rangle\}_{i=0,\dots,d-1}$ with $|0\rangle = |\psi\rangle$. Now $\mathcal{E} \in \mathcal{SI}_b \Rightarrow \mathcal{E} \in \mathcal{MI}_b \Rightarrow \mathcal{E}(|\psi\rangle\langle\psi|) \in \mathcal{I}_b$, thus

$$\mathcal{E}(|\psi\rangle\langle\psi|) = \sum_i q_i |i\rangle\langle i| \tag{4.97}$$

for some probabilities q_i . For $d = 2$, we immediately have $\mathcal{E}(|\psi\rangle\langle\psi|) = p|\psi\rangle\langle\psi| + (1-p)I/2$ for some p which is a function of $|\psi\rangle$. For $d > 2$, take $i, j \neq 0$ such that $i \neq j$. Then there is a rotated basis b' which is equal to b except for the replacement of $|i\rangle, |j\rangle$ by $|\pm_{ij}\rangle := \frac{1}{\sqrt{2}}(|i\rangle \pm |j\rangle)$, such that $|\psi\rangle\langle\psi| \in \mathcal{I}_{b'}$. Now we have

$$\begin{aligned}
\mathcal{E}(|\psi\rangle\langle\psi|) &= \left(\sum_{k \neq i,j} q_k |k\rangle\langle k| \right) \\
&\quad + \frac{1}{2} \left[(q_i + q_j) |+_ {ij}\rangle\langle +_ {ij}| + (q_i + q_j) |-_ {ij}\rangle\langle -_ {ij}| \right. \\
&\quad \left. + (q_i - q_j) |+_ {ij}\rangle\langle -_ {ij}| + (q_i - q_j) |-_ {ij}\rangle\langle +_ {ij}| \right],
\end{aligned} \tag{4.98}$$

so $\mathcal{E}(|\psi\rangle\langle\psi|) \in \mathcal{I}_{b'}$ only if $q_i = q_j$. Since $\mathcal{E} \in \mathcal{SI}_{b'}$, it follows that all q_i for $i > 0$ are equal to some q independent of i .

Therefore we can write

$$\begin{aligned}\mathcal{E}(|\psi\rangle\langle\psi|) &= (1 - q[d - 1])|\psi\rangle\langle\psi| + q \sum_{i>0} |i\rangle\langle i| \\ &= (1 - q[d - 1])|\psi\rangle\langle\psi| + q(I - |\psi\rangle\langle\psi|) \\ &= (1 - qd)|\psi\rangle\langle\psi| + (qd)I/d.\end{aligned}\tag{4.99}$$

Evidently, for any d we can write $\mathcal{E}(|\psi\rangle\langle\psi|) = p|\psi\rangle\langle\psi| + (1 - p)I/d$. However, we must be careful that p is in principle a function of $|\psi\rangle$. The following argument shows that it must in fact be a constant.

Choose an arbitrary basis $b = \{|i\rangle\}_{i=0,\dots,d-1}$. From above, for any i we can write $\mathcal{E}(|i\rangle\langle i|) = p_i|i\rangle\langle i| + (1 - p_i)I/d$. Using Lemma 3,

$$\begin{aligned}I &= \sum_i \mathcal{E}(|i\rangle\langle i|) \\ &= \sum_i p_i|i\rangle\langle i| + (1 - p_i)I/d \\ &= \left(\sum_i p_i|i\rangle\langle i|\right) + \left(1 - \frac{\sum_i p_i}{d}\right)I,\end{aligned}\tag{4.100}$$

which implies $p_i = \text{const.} =: p$. Note that p still depends implicitly on the choice of basis.

For any pair $i \neq j$, construct the state $|\psi\rangle := \sqrt{a}|i\rangle + e^{i\phi}\sqrt{1-a}|j\rangle$ as a function of $a \in [0, 1]$, $\phi \in [0, 2\pi]$. We know from above that

$$\mathcal{E}(|k\rangle\langle k|) = p|k\rangle\langle k| + (1 - p)I/d \quad \forall k,\tag{4.101}$$

$$\mathcal{E}(|\psi\rangle\langle\psi|) = r|\psi\rangle\langle\psi| + (1 - r)I/d,\tag{4.102}$$

where r is implicitly a function of a and ϕ . From the definition of $|\psi\rangle$,

$$\begin{aligned}\mathcal{E}(|\psi\rangle\langle\psi|) &= a \left[p|i\rangle\langle i| + (1 - p)\frac{I}{d} \right] + (1 - a) \left[p|j\rangle\langle j| + (1 - p)\frac{I}{d} \right] \\ &\quad + \sqrt{a(1-a)}\mathcal{E}(e^{-i\phi}|i\rangle\langle j| + e^{i\phi}|j\rangle\langle i|),\end{aligned}\tag{4.103}$$

and from (4.102), we have

$$\mathcal{E}(|\psi\rangle\langle\psi|) = r \left[a|i\rangle\langle i| + (1 - a)|j\rangle\langle j| + \sqrt{a(1-a)}(e^{-i\phi}|i\rangle\langle j| + e^{i\phi}|j\rangle\langle i|) \right] + (1 - r)\frac{I}{d}.\tag{4.104}$$

We equate the two previous expressions and take the $|i\rangle\langle i|$ matrix element. Using the fact that $\mathcal{E} \in \mathcal{ST}_b$ implies $\mathcal{E}(|i\rangle\langle j|)$ is fully off-diagonal, we obtain

$$\begin{aligned} ap + a(1-p)/d + (1-a)(1-p)/d &= ar + (1-r)/d \\ \Rightarrow p\left(a - \frac{1}{d}\right) &= r\left(a - \frac{1}{d}\right) \quad \forall a, \phi. \end{aligned} \quad (4.105)$$

This immediately shows that $r = p \quad \forall \phi, \forall a \neq 1/d$; by continuity of $\mathcal{E}$ we must have equality for all a .

Instead equating the off-diagonal parts of the two expressions for $\mathcal{E}(|\psi\rangle\langle\psi|)$ (again using $\mathcal{E} \in \mathcal{ST}_b$), we have

$$\sqrt{a(1-a)}\mathcal{E}(e^{-i\phi}|i\rangle\langle j| + e^{i\phi}|j\rangle\langle i|) = r\sqrt{a(1-a)}(e^{-i\phi}|i\rangle\langle j| + e^{i\phi}|j\rangle\langle i|) \quad (4.106)$$

for all ϕ . Choosing any $a \neq 0, 1$ and using $r = p$,

$$\phi = 0: \quad \mathcal{E}(|i\rangle\langle j|) + \mathcal{E}(|j\rangle\langle i|) = p(|i\rangle\langle j| + |j\rangle\langle i|) \quad (4.107)$$

$$\phi = \pi/2: \quad -i\mathcal{E}(|i\rangle\langle j|) + i\mathcal{E}(|j\rangle\langle i|) = p(-i|i\rangle\langle j| + i|j\rangle\langle i|), \quad (4.108)$$

so that $\mathcal{E}(|i\rangle\langle j|) = p|i\rangle\langle j|$. Together with $\mathcal{E}(|i\rangle\langle i|) = p|i\rangle\langle i| + (1-p)I/d$, this is all we need to conclude that $\mathcal{E}(\rho) = p\rho + (1-p)I/d$ for any state ρ .

4.4 Conclusions

The aim of this chapter was to connect coherence and discord, two fundamentally quantum resources with historically separate proposals. A strong clue of the existence of such a link is the similarity in the structures of the states for which these quantities vanish. Our first set of results describe the creation of discord when restricted to global incoherent operations; the coherence of the initial state sets a limit on the amount of producible discord, according to a variety of different measures. Moreover, for the global discord (which generalises the original proposal of Ollivier and Zurek to multipartite systems), the creation of discord requires a reduction in coherence of the individual subsystems. Based on this observation, we have given an argument to suggest that the quantum speed-up in the DQC1 algorithm is better signified by the consumption of local coherence than by the creation of discord. Note that a later work by Matera et al. [184] relates the precision of the computed result to a coherence

measure. However, there are still open questions here, in particular relating to the fairest way of comparing the classical and quantum regimes of the algorithm.

Next, we have attempted to better understand the relation between coherence and discord by restricting the latter to be defined with respect to a fixed basis. This is natural given the basis-dependent nature of coherence. Rather than being a mere “special case” of the usual basis-independent discord, we have found that more aspects of quantum correlations are revealed using the link with coherence. In particular, BDD can be interpreted as coherence shared between subsystems. Its behaviour under local incoherent operations is rather subtle – and it may even increase in general under local SI operations. However, unlike the basis-independent case, where local creation of discord is somewhat poorly understood, we have given a physical basis for the increase in BDD under local SI operations. We have shown that such an increase results from the loss of a memory recording classical information about which operation was performed. An alternative view has been given by quantifying the difficulty of recovering a state after complete decoherence, using only local operations. We have shown measures of this type to have the same properties as the BDD quantity δ_b . An interesting question is whether δ_b can itself be directly related to such a recovery task.

Finally, we have provided a new connection between the free operations for coherence and the full set of local channels which cannot create discord. Namely, a large subset of the latter channels is unitarily equivalent to those channels which are SI with respect to every basis. We hope that this result will help to inform the correct choice of free operations for a potential resource theory of discord. It might also be combined with the results on the behaviour of BDD under local operations in order to derive new results about the behaviour of basis-independent discord. For example, it is plausible that $\delta(B|A)$ is monotonic under depolarising channels on A , but such a statement does not exist in the literature.

Chapter 5

Quantifying macroscopic coherence

This chapter identifies the type of coherence resource theory necessary for distinguishing microscopic from macroscopic superpositions. Starting from a set of axioms, we show that this resource theory is equivalent to a particular instance of the resource theory of asymmetry. Our main proposal is that quantum macroscopicity should be quantified in terms of measures that are valid monotones in this resource theory. To this end, we identify a particular measure given by the quantum Fisher information, showing that it can be interpreted as an asymptotic formation rate. We then evaluate some previous proposals for measures of macroscopicity, finding connections with the Fisher information quantity, and pointing out deficiencies such as lack of monotonicity.

5.1 Defining the problem

In quantifying speakable coherence in the previous chapters, we have assumed preferred basis states whose mixtures define the incoherent states. Let us now add some additional structure and consider a preferred *observable* A . For any given physical situation to be analysed, we assume that such an observable is one which is justifiably used for the macroscopic scale description of the system. This may seem a little ill-defined, but an observable of this type is easily found (although perhaps not uniquely) for the types of system typically considered: for example, the total momentum of a large diffracting molecule, a field quadrature for a mode of light, or the current in a superconducting circuit. Importantly, we must agree that a superposition of states that are sufficiently distinguishable in a measurement of this observable is

to be considered non-classical.

By its spectral decomposition $A = \sum_i a_i |i\rangle\langle i|$ – where we make no assumptions about degeneracy, so two a_i may be equal – this also identifies a preferred basis, but now each $|i\rangle$ has an associated real eigenvalue a_i . Throughout this section we use discrete sums, but integrals are necessary for continuous spectra. This lets us attribute a scale to a superposition of basis states. We say that a superposition of the form $|i\rangle + |j\rangle$ contains δ -coherence, where $\delta = a_i - a_j$. Hence we are able to distinguish macroscopic and microscopic superpositions by their values of δ . For a general state, we define:

Definition 18. *With respect to a preferred observable A , the δ -coherence of a state $\rho = \sum_{i,j} \rho_{ij} |i\rangle\langle j|$ is*

$$\rho^{(\delta)} := \sum_{(i,j) \sim \delta} \rho_{ij} |i\rangle\langle j|, \quad (5.1)$$

where the notation $(i,j) \sim \delta$ means i,j such that $a_i - a_j = \delta$.

The state can be expressed as the sum of its parts: $\rho = \sum_{\delta \in \Omega} \rho^{(\delta)}$, where $\Omega := \{a_i - a_j\}_{i,j}$ is the set of all eigenvalue differences of A . Also note that $\rho^{(\delta)\dagger} = \rho^{(-\delta)}$, so it is not important to be careful about the sign of δ when discussing its size. In addition, any two elements are orthogonal according to the Hilbert-Schmidt inner product: $\text{Tr}[\rho^{(\delta)} \sigma^{(\delta')}] = 0$ for any ρ, σ and $\delta \neq \delta'$.

The set of free states in the resource theory is exactly the same as for speakable coherence: incoherent mixtures of the basis states $|i\rangle$. Equivalently, a state is free iff $\rho = \rho^{(0)}$. The only difference is that now, we allow for coherence within degenerate subspaces.

We could imagine quantifying the amount of δ -coherence via $\|\rho^{(\delta)}\|$ with some suitable norm. Each state will have its own signature distribution of $\|\rho^{(\delta)}\|$ as a function of δ ; for instance, a cat state is expected to have a peak around a large value of δ . Therefore we consider δ -coherence for each δ to all be quantities of independent interest. In order to keep this structure consistent, we must impose that the free operations cannot turn δ -coherence into δ' -coherence with $\delta \neq \delta'$. Without this constraint, non-zero δ' -coherence could be created from a state containing only δ -coherence. So a free operation $\mathcal{E}$ must be such that

$$\mathcal{E}(\rho)^{(\delta)} = \mathcal{E}(\rho^{(\delta)}) \quad \forall \delta \in \Omega. \quad (5.2)$$

This is the only absolutely necessary assumption on the free operations, and we shall make no further assumptions. Note that the condition (5.2) for $\delta = 0$ is the same as saying that $\mathcal{E}$ is incoherent – either in the MI sense for channels, or in the BCP sense if applied to individual Kraus operators.

Just as for speakable coherence, we can give conditions to be satisfied by any valid measure C of macroscopic coherence:

- (M1) $C(\rho) = 0$ iff $\rho \in \mathcal{I}$, or equivalently, $\rho = \rho^{(0)}$;
- (M2a) C is a deterministic monotone under free operations;
- (M2b) C is an ensemble monotone under free operations;
- (M3) C is convex.

Another condition can be given in order to capture the requirement that a more macroscopic superposition be given a higher value of macroscopic coherence. The following is one possible way to state this, by imposing a strict order on the simplest superpositions:

- (M4) Let $|\psi\rangle = (|i\rangle + |j\rangle)/\sqrt{2}$, $|\phi\rangle = (|k\rangle + |l\rangle)/\sqrt{2}$. If $|a_i - a_j| > |a_k - a_l|$ then $C(|\psi\rangle\langle\psi|) > C(|\phi\rangle\langle\phi|)$.

We must also define how to combine systems. Given two systems with preferred observables A_1 and A_2 , what is the corresponding observable for both systems together? A natural choice is to take the sum $A_{12} = A_1 \otimes I + I \otimes A_2$. This can be justified by noting that observables deemed to be macroscopic are often additive over different subsystems. Therefore this choice of A_{12} uniquely selects one member of the possible global additive observables reducing to A_1, A_2 on the subsystems.

5.2 Digression: asymmetry and reference frames

We have now built up the required mathematical structure for the resource theory needed to describe macroscopic coherence. It turns out that this is equivalent to a resource theory that is already known: that of asymmetry. This was developed to describe situations in which operations are constrained by symmetries and conservation laws [101, 115, 185–187]. In its full generality, a resource theory of asymmetry

is defined with respect to a group G whose elements correspond to unitary transformations of quantum states. The free states and operations are symmetric with respect to the group action, so asymmetric objects constitute resources. Reference frames are such objects – for example, a set of Cartesian axes in space breaks the rotational symmetry group $SO(3)$, and allows for asymmetric operations to be performed with the axes as reference. This formalism has been applied to discussions of quantum communication protocols between parties who lack full information about each other’s reference frames [188, 189]. In order to communicate, say, a direction in space without any shared reference frame, one has to send a physical system; this kind of information has been termed “unspeakable” [114].

In order to make a connection with coherence, note that the unitary transformations generated by the observable A , i.e.

$$\mathcal{U}_t(\rho) := e^{-itA}\rho e^{itA}, \quad t \in \mathbb{R}, \quad (5.3)$$

form a representation of either the group $U(1)$ or the real line under addition $(\mathbb{R}, +)$. The typical case is $(\mathbb{R}, +)$, but $U(1)$ is found for example whenever the eigenvalues a_i are all integer multiples of some constant, as the transformation is then periodic. In this context, the δ -coherence structure (5.1) is known as a decomposition into **modes of asymmetry** [187]. It is easily seen that each term acquires a uniformly rotating phase under the group action:

$$\mathcal{U}_t(\rho^{(\delta)}) = e^{-it\delta} \rho^{(\delta)}. \quad (5.4)$$

For this reason, one can think of the decomposition as a kind of Fourier analysis.

The free states of the resource theory are symmetric with respect to these transformations, satisfying $\mathcal{U}_t(\rho) = \rho \forall t$. By (5.4), this is equivalent to $\rho = \rho^{(0)}$, as well as $\rho \in \mathcal{I}$. The free operations, known as **translationally invariant**, are defined by the following condition (recalling Definition 8): $[\mathcal{E}, \mathcal{U}_t] = 0 \forall t$. As originally shown in [187], this is equivalent to our condition (5.2) on free operations for macroscopic coherence. To see this, write

$$\mathcal{E} \circ \mathcal{U}_t(\rho) = \sum_{\delta} e^{-it\delta} \mathcal{E}(\rho^{(\delta)}), \quad (5.5)$$

$$\mathcal{U}_t \circ \mathcal{E}(\rho) = \sum_{\delta} e^{-it\delta} [\mathcal{E}(\rho)]^{(\delta)}. \quad (5.6)$$

By the linear independence of the functions $e^{-it\delta}$, we have $[\mathcal{E}, \mathcal{U}_t] = 0$ iff these two expressions are equal term-wise.

Yet another equivalent statement can be made in terms of Kraus operators – recall the claim in Chapter 3 that an operation is TI iff it has a set of Kraus operators satisfying $K_\mu = K_\mu^{(\delta_\mu)}$ (using the presently introduced notation). It is easily verified that a single such Kraus operator preserves the structure of δ -coherence – i.e., the operation $\mathcal{E}_\mu(\rho) = K_\mu \rho K_\mu^\dagger$ satisfies (5.2). This is clearly also true for any mixture of these Kraus operators. The converse is less simple, and relies on Lemma 1 of [101]. This Lemma applies to any group; it says that a set of Kraus operators can be found for any TI operations, such the group action applied to the Kraus operators simply mixes them according to a unitary matrix irreducible representation of G . For $U(1)$ (as in Equation 41 of [101]) or $(\mathbb{R}, +)$, this becomes $\mathcal{U}_t(K_\mu) = e^{-it\delta_\mu} K_\mu$, as the functions $e^{-it\delta_\mu}$ are the only unitary irreducible representations of these groups.

A connected result gives the full physical relevance of TI operations in the theory of asymmetry. This is the dilation theorem described in (3.35) of Chapter 3 – any TI operation on a system can be constructed with an ancilla, using a global unitary and a measurement on the ancilla, both of which are TI with respect to a global additive observable.

The theory of $U(1)$ asymmetry has been successfully used to clarify accounts of phase coherence, notably in quantum optics and condensates [190–193]. Here, a state which is sensitive to the transformations $\mathcal{U}_t$ is said to act as a phase reference (where the phase is conjugate to the generating observable A). For these reasons, $U(1)$ asymmetry was suggested as a way to quantify coherence in [187], as a more physically motivated alternative to the BCP proposal. It was only later that the notions of speakable and unspeakable coherence were disentangled [41].

5.3 Monotones

5.3.1 Distance-based measures

We can now employ the results and methods of the resource theory of asymmetry to find valid measures satisfying our criteria. Firstly, we of course find that any monotone under the largest set of incoherent operations, namely MI, satisfies (M1) and (M2a). This applies to any distance-based measure of the form $C_D(\rho) = \min_{\sigma \in \mathcal{I}} D(\rho, \sigma)$. For

example, the relative entropy of coherence existed in earlier works under the name of the relative entropy of asymmetry (or of “frameness”) [194].

Since the TI operations are also a subset of the SI operations, the measures proposed earlier, $C'_D(\rho) = D(\rho, \Phi(\rho))$, are also valid. Note that the dephasing operation has to be written a little differently in order to take into account the possible degeneracy in the spectrum of A – so $\Phi(\rho) = \sum_a \Pi_a \rho \Pi_a$, where the sum is over distinct eigenvalues of A and each Π_a is the projector onto the corresponding eigenspace.

While these measures satisfy (M1), (M2a) and, in certain cases (such as the relative entropy) (M2b) and (M3), they always fail (M4). Indeed, it is a requirement of the theory of speakable coherence that any two superpositions of the form $(|i\rangle + |j\rangle)/\sqrt{2}$ have the same value for any measure. This is because they are related by a permutation of the incoherent basis, which is a free unitary. Therefore we need a different approach to construct measures of macroscopic coherence (see also Chapter 6 for another alternative).

5.3.2 Variance and QFI

It was shown in Lemma 8 of [101] that the variance $V(\rho, A) = \text{Tr}[\rho A^2] - (\text{Tr}[\rho A])^2$ of A with respect to ρ satisfies (M2a,b) in the $U(1)$ case, as long as we are only concerned with pure states (for monotonicity, this means that both the input and output states must be pure). We do not repeat their proof here, but note that it also extends to the $(\mathbb{R}, +)$ case. The only way in which the proof could differ is in the evaluation of the commutator $[A, K_\mu]$. In that work, the observable $A = \sum_{n=0}^{\infty} n |n\rangle\langle n|$ has a harmonic spectrum, and then $[A, K_\mu] = k_\mu K_\mu$ for some integer k_μ . For general A , we know that we can take $K_\mu = K_\mu^{(\delta_\mu)}$. It is then easily verified that $[A, K_\mu] = \delta_\mu K_\mu$, and the rest of the proof continues unaffected by the replacement of k_μ by δ_μ . It is trivial to see that (M1) (for pure states) and (M4) are satisfied.

To extend this to mixed states, we need a measure which reduces to the variance for pure states. This is achieved by the quantum Fisher information of the family of states $\rho_t = \mathcal{U}_t(\rho)$ with the parameter t . Recall from Equation (2.45) that the QFI at some parameter value t_0 is $\mathcal{F}(\rho_t)|_{t_0} = 4\partial_t^2 D_B^2(\rho_{t_0}, \rho_t)|_{t_0}$. This can be expressed as [90]

$$\mathcal{F}(\rho, A) = 2 \sum_{i,j} \frac{(\lambda_i - \lambda_j)^2}{\lambda_i + \lambda_j} |\langle \psi_i | A | \psi_j \rangle|^2, \quad (5.7)$$

where $\rho = \sum_i \lambda_i |\psi_i\rangle\langle\psi_i|$ is a spectral decomposition and the sum neglects terms for which $\lambda_i = \lambda_j = 0$. Note that this is in fact independent of t_0 , so we omit it. From now on, we refer to $\mathcal{F}(\rho, A)$ as the QFI of ρ with respect to A .

The QFI is always bounded from above by the variance up to a constant factor, $\mathcal{F}(\rho, A) \leq 4V(\rho, A)$, and equality is achieved in the case of a pure state. For general states, we have the following:

Result 14. *The QFI satisfies the criteria (M1-4).*

Proof. (M1) can be seen from the expression (5.7) – each term in the sum is non-negative, and vanishes iff $\langle\psi_i|A|\psi_j\rangle = 0$ or $\lambda_i = \lambda_j$. For $\lambda_i \neq \lambda_j$, the off-diagonal matrix element $\langle\psi_i|A|\psi_j\rangle$ vanishes. In degenerate subspaces, there exists a change of eigenbasis such that these elements can be made to vanish. Hence, when all terms vanish, A and ρ are simultaneously diagonalisable, so $\rho \in \mathcal{I}$.

For (M2-3), we recall a concept known as a **convex roof** construction. This is a way of extending a real-valued function f defined for pure states, to mixed states. A general state can be expressed in many different ways as an ensemble of pure states, $\rho = \sum_\alpha p_\alpha |\psi_\alpha\rangle\langle\psi_\alpha|$, where the $|\psi_\alpha\rangle$ need not be orthogonal. The convex roof of f is given by a minimisation over all possible ensemble decompositions of ρ :

$$f_{\text{CR}}(\rho) := \min_{\{p_\alpha, |\psi_\alpha\rangle\}} \sum_\alpha p_\alpha f(|\psi_\alpha\rangle). \quad (5.8)$$

It was conjectured [195] and later proven [196] that $\frac{1}{4}\mathcal{F}(\rho, A)$ is the convex roof of $V(|\psi\rangle\langle\psi|, A)$. The convexity of the QFI [195] follows directly from this statement.

Suppose that a TI operation $\mathcal{E}$ with Kraus operators K_μ is performed on an initial state ρ . Choose the optimal ensemble $\{p_\alpha, |\psi_\alpha\rangle\}$ in (5.8), so that $\mathcal{F}(\rho, A) = 4 \sum_\alpha p_\alpha V(|\psi_\alpha\rangle\langle\psi_\alpha|, A)$. Also define $\sqrt{p_{\mu|\alpha}} |\phi_{\mu|\alpha}\rangle = K_\mu |\psi_\alpha\rangle$, with $p_{\mu|\alpha} = |K_\mu |\psi_\alpha\rangle|^2$. Using the monotonicity of the variance, we have

$$\begin{aligned} \mathcal{F}(\rho, A) &\geq 4 \sum_\alpha p_\alpha \sum_\mu p_{\mu|\alpha} V(|\phi_{\mu|\alpha}\rangle\langle\phi_{\mu|\alpha}|, A) \\ &= \sum_{\alpha, \mu} p_\alpha p_{\mu|\alpha} \mathcal{F}(|\phi_{\mu|\alpha}\rangle\langle\phi_{\mu|\alpha}|, A). \end{aligned} \quad (5.9)$$

Let the output states be $q_\mu \sigma_\mu = K_\mu \rho K_\mu^\dagger$, then we have

$$q_\mu \sigma_\mu = \sum_\alpha p_\alpha K_\mu |\psi_\alpha\rangle\langle\psi_\alpha| K_\mu^\dagger = \sum_\alpha p_\alpha p_{\mu|\alpha} |\phi_{\mu|\alpha}\rangle\langle\phi_{\mu|\alpha}|. \quad (5.10)$$

Together with the convexity of the QFI, it follows that $\mathcal{F}(\rho, A) \geq \sum_{\mu} q_{\mu} \mathcal{F}(\sigma_{\mu}, A)$, proving (M2b). Finally, (M2a) results from applying convexity once more to give $\sum_{\mu} q_{\mu} \mathcal{F}(\sigma_{\mu}, A) \geq \mathcal{F}(\mathcal{E}(\rho), A)$. $\square$

5.3.3 Generalised QFI

In fact, it turns out that the QFI is not uniquely determined by the requirement of being a monotone and reducing to the variance for pure states. A whole class of such measures can be constructed. Since this is only tangential to the discussion of macroscopicity, we just give an outline here and leave the details to Appendix A. Each measure is associated with a function f (chosen from a particular class) on the real numbers, and can be written as

$$I_f(\rho, A) = \frac{f(0)}{2} \| -i[A, \rho] \|_{\rho, f}^2, \quad (5.11)$$

where $\|\cdot\|_{\rho, f}$ is a norm which depends on both f and ρ . The point is that I_f measures the size of the derivative $-i[A, \rho] = \partial_t \mathcal{U}_t(\rho)$ using a norm with the special property of being monotonic under arbitrary channels, i.e. $\|\mathcal{E}(X)\|_{\mathcal{E}(\rho), f} \leq \|X\|_{\rho, f}$.

These metrics and generalised QFI quantities were first studied by Morozova and Centsov [197] and later by Petz [198, 199]. Whenever $f(0) \neq 0$ (a so-called regular case), we have $I_f(|\psi\rangle\langle\psi|, A) = V(|\psi\rangle\langle\psi|, A)$. $\frac{1}{4}\mathcal{F}$ is the largest of these measures – as it must be, given that the convex roof of the variance is the largest convex function reducing to the variance for pure states [195]. Another notable element is the Wigner-Yanase-Dyson skew information [200], defined as $I_{sk}(\rho, A) := -\frac{1}{2} \text{Tr}([A, \sqrt{\rho}]^2)$. This was proven to be an asymmetry monotone in [201], and is always within a constant factor of the QFI, $I_{sk}(\rho, A) \leq \frac{1}{4}\mathcal{F}(\rho, A) \leq 2I_{sk}(\rho, A)$ [202].

5.4 QFI as an asymptotic preparation cost

In this section, we describe how the QFI plays a distinguished role in the present resource theory: it quantifies the cost of preparing a general state as a mixture of pure coherence “units”, in the asymptotic limit. Firstly we look at pure states, building on the results of [101, 203]. In particular, the latter studied the following question: when can a large number $N \gg 1$ of copies of $|\psi\rangle$ be transformed under TI operations into $M \gg 1$ copies of another state $|\phi\rangle$, and what is the achievable N/M

in the asymptotic limit $N, M \rightarrow \infty$? The ratio N/M turns out to equal the ratio of the variances. However, different subtleties are encountered according to the nature of the spectrum of the observable A ; we treat the cases separately. In accordance with the earlier remarks, given N independent copies $|\psi\rangle^{\otimes N}$, each one is associated with its own copy of the observable A . Then the global free operations are defined with respect to the additive observable $\sum_{i=1}^N A_i$, where A_i acts only on the i th copy.

5.4.1 Continuous spectrum

The simplest case is that of a continuous variable $A = \int_{-\infty}^{\infty} dx x \Pi_x$. The global observable can be written as $\sum_{i=1}^N A_i = \int dx x \Pi_{N,x}$, where $\Pi_{N,x}$ projects onto the eigenspace $S_{N,x}$ which contains all states on $\mathcal{H}^{\otimes N}$ with a total eigenvalue x . Then we can write

$$|\psi_N\rangle := |\psi\rangle^{\otimes N} = \int dx \sqrt{p_N(x)} |\Psi_{N,x}\rangle, \quad (5.12)$$

where $|\Psi_{N,x}\rangle \in S_{N,x}$ and the coefficients can be assumed to be real and positive without loss of generality. The same may be done for the final state, i.e. $|\phi_M\rangle := |\phi\rangle^{\otimes M} = \int dx \sqrt{q_M(x)} |\Phi_{M,x}\rangle$. If $N < M$, then in order to keep the input and output spaces the same, we use this notation as shorthand for $|\psi\rangle^{\otimes N} \otimes |0\rangle^{M-N}$, where $|0\rangle$ is some free pure state with zero eigenvalue. The same remarks apply for $|\phi\rangle^{\otimes M}$ if $N > M$.

The main idea of the proof is the central limit theorem in statistics. Loosely speaking, $p_N(x)$, as the probability density function for a sum of N independent and identically distributed random variables, tends to a Gaussian distribution as N increases. The same applies to $q_M(x)$, and we would like to make the two distributions match up by controlling their means and variances – the only two parameters needed to uniquely specify a Gaussian. Denoting the mean and variance of $|\psi\rangle$ by $\mu_\psi = \langle \psi | A | \psi \rangle$, $V_\psi = V(|\psi\rangle\langle\psi|, A)$, the corresponding parameters for $p_N(x)$ are $N\mu_\psi$, NV_ψ (and similarly for $|\phi\rangle$).

Firstly, we set the variances of $p_N(x)$ and $q_M(x)$ to be equal; this requires $N/M = V_\phi/V_\psi$. For finite numbers of copies this is not in general possible, but can be approximated arbitrarily well (with an error of size $O(1/M)$) by taking sufficiently large N, M . Next, to make the means equal, we will shift $p_N(x)$ by $s := M\mu_\phi - N\mu_\psi$. This is accomplished by the transformation $|\Psi_{N,x}\rangle \rightarrow |\Phi_{M,x+s}\rangle \forall x$, which can be done deterministically by a free operation. Therefore we have the overall transformation

$$|\psi\rangle^{\otimes N} \rightarrow |\psi'_N\rangle = \int dx \sqrt{p_N(x)} |\Phi_{N,x+s}\rangle. \quad (5.13)$$

The fidelity with the desired state is $\langle \psi'_N | \phi_N \rangle = \int dx \sqrt{p_N(x-s)q_M(x)}$, which is also the classical fidelity between the distributions $p_N(x-s)$ and $q_M(x)$. For sufficiently large N and M , each of these is arbitrarily close to a Gaussian distribution of the same mean and variance. The two Gaussians have the same mean; the $O(1)$ difference in the variances is much less than their $O(N)$ size. Hence the fidelity between these Gaussians approaches 1 in the asymptotic limit, as does the fidelity between $|\psi'_N\rangle$ and $|\phi_M\rangle$.

5.4.2 Harmonic spectrum

Now suppose that A has an equally spaced (or “harmonic”) spectrum; for simplicity, we assume that it is integer-valued, as the arguments are unchanged under an affine transformation. So we write $A = \sum_n n \Pi_n$ and $\sum_{i=1}^N A_i = \sum_n n \Pi_{N,n}$, where the range of values may be infinite or bounded. The states are expressed in the form $|\psi_N\rangle = |\psi\rangle^{\otimes N} = \sum_n \sqrt{p_N(n)} |\Psi_{N,n}\rangle$ as before.

Compared with the continuous case, there are some additional subtleties. Firstly, the mean of $p_N(n)$ can only be shifted by an integer, so cannot in general exactly match the mean of $q_M(n)$. However, by shifting by an amount $s := [M\mu_\phi - N\mu_\psi]$ (brackets denoting the nearest integer), the error is only $O(1)$. Again, this error is negligible compared with the $O(N)$ size of the mean, so presents no obstacle to the fidelity of the states approaching unity. In addition, these shifts may not be completely deterministic – if some of the values are trying to be moved outside of the spectrum of $\sum_{i=1}^N A_i$. Nevertheless, the probability of failure is negligible, as the tails of the Gaussian distribution are exponentially small for large N .

A more difficult problem is encountered when the distribution of n in a state contains gaps. For this reason, the proof in [101] only stated the result for pure states with “gapless number spectra”, meaning that $p(n) > 0$ in some (possibly infinite) interval, with $p(n) = 0$ outside this interval. For example, a distribution with support on $\{2, 3, 4, 5\}$ is gapless, while one on $\{2, 4, 5\}$ is not.

However, with the help of a more powerful version of the central limit theorem, we can overcome this and give a more general statement. For this, we need to introduce some terminology: let the **span** of an integer-valued distribution $p(n)$ be the largest natural number h such that the support of $p(n)$ is contained in $h\mathbb{Z} = \{0, \pm h, \pm 2h, \dots\}$.

For example, a distribution with support $\{2, 4, 6\}$ has a span of 2.

A version of the central limit theorem exists for integer-valued distributions [204] (Chapter 7, Theorem 1), and guarantees that for $p(n)$ with span h ,

$$p_N(n) \rightarrow \frac{h}{\sqrt{2\pi N}\sigma} e^{-\frac{(nh-N\mu)^2}{2N\sigma^2}} + o(N^{-1/2}) \quad \text{as } N \rightarrow \infty. \quad (5.14)$$

In words, $p_N(n)$ tends to a distribution of the same span h with a Gaussian envelope.

Now it is easy to see that the proof used for the continuous case still holds here, as long as $p(n)$ and $q(n)$ have the same span – whether or not they are gapless. What if their spans are different? As noted in [101], the conversion rate may be zero. For example, if $|\psi\rangle = (|0\rangle + |2\rangle)/\sqrt{2}$ and $|\phi\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$, then $|\psi\rangle^{\otimes N} \rightarrow |\phi\rangle^{\otimes M}$ can never be achieved – to do so would require the creation of new δ -coherence (with $\delta = 1$). However, the reverse transformation is actually possible. We now prove this observation in the general case.

Let $p(n)$ have span 1, and $q(n)$ have span $h > 1$. Then the asymptotic rate of conversion $|\psi\rangle^{\otimes N} \rightarrow |\phi\rangle^{\otimes M}$ is still $N/M = V_\phi/V_\psi$. To show this, note that the means and variances can still be made approximately equal, as before. However, p_N and q_M are now discrete distributions with the same Gaussian envelope but with different spans (1 and h , respectively).

Suppose we perform a measurement on $|\psi'_N\rangle$ which projects onto the subspaces $\mathcal{L}_k$, $k = 0, 1, \dots, h-1$, spanned by the eigenstates of $\sum_{i=1}^N A_i$ with eigenvalues in $h\mathbb{Z} + k$. For example, with $h = 2$, $\mathcal{L}_0$ contains even numbers and $\mathcal{L}_1$ contains odd numbers. This measurement is a free operation. When the outcome corresponding to $\mathcal{L}_k$ is obtained, we shift the resulting state by $-k$, i.e. $|\Phi_{N,n}\rangle \rightarrow |\Phi_{N,n-k}\rangle$. So each outcome k gives rise to a state with a distribution having the same support as q_M , and a Gaussian envelope very close to that of q_M . Note that the shift of $-k$ is $O(1)$ and thus negligible compared with the mean of $O(N)$ in the asymptotic limit. Hence we have deterministically produced $|\phi\rangle^{\otimes M}$ to arbitrary accuracy.

5.4.3 Arbitrary discrete spectrum

Finally, we have the case of $A = \sum_a a\Pi_a$, with a discrete spectrum containing arbitrary values. Unfortunately, it seems that the above proofs no longer work. On one hand, this is not usually an issue since most of the examples that we encounter fit into the previous two cases. On the other, we would argue that due to the finite precision of

real measurements, it is physically reasonable to replace A by another observable $\tilde{A}$ whose eigenvalues are contained in $\epsilon\mathbb{Z}$, mapping each eigenvalue of A onto the closest eigenvalue of $\tilde{A}$. Here, ϵ must be both below conceivable experimental resolution and small enough that the eigenvalues are mapped one-to-one. We can then treat A as $\tilde{A}$ for all practical purposes.

5.4.4 Formation of mixed states

To summarise the above results for pure states:

Result 15. *There exists a **coherence unit** $|\psi_0\rangle$ such that any pure state $|\psi\rangle$ can be prepared in the asymptotic limit deterministically by the TI transformation $|\psi_0\rangle^{\otimes N} \rightarrow |\psi\rangle^{\otimes M}$, at a rate*

$$\lim_{N \rightarrow \infty} \frac{N}{M} = \frac{V(|\psi\rangle\langle\psi|, A)}{V(|\psi_0\rangle\langle\psi_0|, A)}. \quad (5.15)$$

We can generally choose $|\psi_0\rangle$ to be any coherent state. However, when A has a harmonic spectrum of the form $a_n = a_0 + bn$, $n \in \mathbb{Z}$, we must ensure that there exists an n such that the amplitudes for a_n and a_{n+1} are nonzero.

An operational meaning can be given to the QFI by considering the preparation of mixed states [205]. Suppose we decompose an arbitrary state as an ensemble of pure states, $\rho = \sum_{\mu} p_{\mu} |\psi_{\mu}\rangle\langle\psi_{\mu}|$. One can then create $\rho^{\otimes M}$ from $|\psi_0\rangle^{\otimes N}$ by preparing Mp_{μ} copies of each $|\psi_{\mu}\rangle$, and then forgetting the value of the index μ . The total cost is $N = \sum_{\mu} Mp_{\mu} \times (\text{cost of } |\psi_{\mu}\rangle)$. By choosing the least costly ensemble preparation, we have

$$\begin{aligned} \lim_{N \rightarrow \infty} \frac{N}{M} &= \min_{\{p_{\mu}, |\psi_{\mu}\rangle\}} p_{\mu} \frac{V(|\psi_{\mu}\rangle\langle\psi_{\mu}|, A)}{V(|\psi_0\rangle\langle\psi_0|, A)} \\ &= \frac{\mathcal{F}(\rho, A)}{4V(|\psi_0\rangle\langle\psi_0|, A)}, \end{aligned} \quad (5.16)$$

using the fact that the QFI is the convex roof of the variance. Hence $\frac{1}{4}\mathcal{F}(\rho, A)$ is the $U(1)$ or $(\mathbb{R}, +)$ **asymmetry of formation** (as named in [206], without a link to the QFI).

Note that this is not the same as the **asymmetry cost**, which would have to be defined as the optimal rate for $|\psi_0\rangle\langle\psi_0|^{\otimes N} \rightarrow \rho^{\otimes M}$. This might have a lower value, as the prepare-and-mix strategy may not be the best one. In addition, we do not have

a result about the distillation of asymmetry. Even in the pure state case this is non-trivial: recall from the harmonic spectrum discussion that, for instance, $(|0\rangle + |2\rangle)/\sqrt{2}$ cannot be distilled into $(|0\rangle + |1\rangle)/\sqrt{2}$. While a single coherence unit may be sufficient to prepare arbitrary states, distillation has to result in a set of different units.

5.5 From coherence to macroscopicity

We now discuss how the above concepts and results may be applied to assess quantum macroscopicity. We evaluate two prominent measures proposed in recent years, and analyse how they fit into the present resource-theoretic framework.

5.5.1 Measure by Fröwis and Dür

The authors of [69] studied the setting of N qubits, defining a new measure of macroscopicity and comparing it against other proposals in the literature. The relevant set $\mathcal{A}$ of macroscopic observables in this situation is chosen to be the set of $A = \sum_{i=1}^N A_i$ such that each A_i acts only on qubit i and has eigenvalues ± 1 (so can be written in terms of Pauli operators as $A_i = \mathbf{n}_i \cdot \boldsymbol{\sigma}_i = n_i^x \sigma_i^x + n_i^y \sigma_i^y + n_i^z \sigma_i^z$, with $|\mathbf{n}_i| = 1$). Thinking of each qubit as a spin-1/2 system, such observables could be described as total spins. The measure proposed in [69] is

$$N_F(\rho) := \max_{A \in \mathcal{A}} \frac{\mathcal{F}(\rho, A)}{4N}. \quad (5.17)$$

Their motivation came from earlier proposals [67, 207, 208] to regard the variance $V(|\psi\rangle\langle\psi|, A)$ as a signature of macroscopic quantumness in the case of pure states. The idea there was to look at the scaling of the variance with N , for any given family of states defined as a function of N . On one hand, a product state has $V \propto N$, whereas a macroscopic superposition such as a GHZ state can be expected to have $V \propto N^2$, showing macroscopic quantum fluctuations in the value of the observable A .

Fröwis and Dür then suggested the QFI as an extension to mixed states – mainly based on its ability to distinguish a GHZ state $(|0\rangle^{\otimes N} + |1\rangle^{\otimes N})/\sqrt{2}$ from the corresponding mixture $(|0\rangle\langle 0|^{\otimes N} + |1\rangle\langle 1|^{\otimes N})/2$. Other motivating factors were its ability to detect multipartite entanglement (to be discussed in detail in Chapter 7) and the quantum advantage as described by the Heisenberg limit. The normalisation factor

of $4N$ ensures that $N_F \in [0, N]$, so that it might be regarded in some sense as a “quantum size”.

From our perspective, the measure N_F can be justified as quantifying the maximal coherence with respect to any macroscopic observable. It is sensible to take the maximum; this means that a state is described as macroscopically quantum when there exists *some* macroscopic observable for which the coherence is large. Here, the qualifier “large” is defined as being relative to the size N – we will just accept this for now, and further discuss its meaning in Chapter 7. Note that, by choosing the eigenvalues of A_i to be ± 1 , the spectrum of every $A \in \mathcal{A}$ is harmonic, so the results of Section 5.4 apply.

In order to provide some intuition for N_F , we summarise some known results for certain important classes of pure states. Firstly, any product state has $N_F = 1$, i.e. minimal macroscopicity. This can be seen by writing the state as $|0\rangle^{\otimes N}$ without loss of generality, since N_F is clearly invariant under local unitaries. One choice of optimal observable is $X := \sum_{i=1}^N \sigma_i^x$. At the other extreme, the GHZ state has maximal macroscopicity $N_F = N$, the optimal observable being $Z := \sum_{i=1}^N \sigma_i^z$. Finally, in between these limits we have the family of **Dicke states**

$$|N, k\rangle := \binom{N}{k}^{-\frac{1}{2}} \sum_{\pi} \mathcal{P}_{\pi} |0\rangle^{\otimes(N-k)} |1\rangle^{\otimes k}, \quad (5.18)$$

where $\mathcal{P}_{\pi}$ rearranges the subsystems by the permutation π ; the sum is over all permutations. We have $N_F = 1 + 2f(1 - f)N$, where $f = k/N$, with optimal observable X [69]. $f = 0, 1$ corresponds to a product state, whereas $f = 1/2$ gives $N_F \approx N/2$ for $N \gg 1$.

It was later noted that this measure can also be applied to CV systems [209, 210]. One now has N bosonic modes, each with annihilation and creation operators $a_i, a_i^{\dagger}$. The definition of N_F is the same as (5.17), but now the set of macroscopic variables is chosen to contain sums of quadratures:

$$\begin{aligned} x^{\theta} &:= \sum_{i=1}^N x_i^{\theta_i} \\ x_i^{\theta_i} &:= \cos(\theta_i) x_i + \sin(\theta_i) p_i. \end{aligned} \quad (5.19)$$

We can justify using quadratures by thinking about the classical limit – it could be argued that the classical field description of light is most naturally done in terms

of quadratures. For mechanical oscillators, the classical description is certainly in terms of position and momentum. If we instead calculated coherence with respect to the number operator, then large-number Fock states – typically considered highly non-classical and hard to produce experimentally – would be treated as free.

Since each subsystem is itself an infinite-dimensional system, unlike the case of qubits, macroscopic quantum behaviour can even be found in the case $N = 1$. A number of interesting observations can be made here. Focusing first on a fixed quadrature x , the free operations with respect to x consist of combinations of two elementary kinds of operations:

$$\begin{aligned} \text{displacements} \quad & \rho \rightarrow D(\alpha)\rho D(\alpha)^\dagger, \quad D(\alpha) = e^{\alpha a^\dagger - \alpha^* a}, \\ \text{generalised quadrature measurements} \quad & \rho \rightarrow M\rho M^\dagger, \quad M = \int dx f(x) |x\rangle\langle x|. \end{aligned} \quad (5.20)$$

This can be seen from our earlier result that a free operation has Kraus operators which shift the spectrum of x along with arbitrary amplitude modification. Displacements are in fact free with respect to *every* quadrature. The same cannot be said for the second set of operations – for example, a very fine-grained measurement of x will project the state onto one with large variance in p .

The pure states with minimal N_F are the coherent states $|\alpha\rangle$. To see this, the uncertainty principle (2.6) for any pair of conjugate quadratures x, p says that $V_x V_p \geq 1/4$, where V_x, V_p are the corresponding variances. Then $\max\{V_x, V_p\} \geq \sqrt{V_x V_p} \geq 1/2$, so the smallest maximum is achieved when $V_x = V_p = 1/2$. This also saturates the uncertainty principle; it can be shown that the only pure states with this property are Gaussian [211]. Now the Gaussian states with variance of $1/2$ in every quadrature are the coherent states.

For the squeezed vacuum state, the product of variances along the principal axes (maximum and minimum variances) is still $V_x V_p = 1/4$. From the relation $x^2 + p^2 = 2\hat{n} + 1$ and the fact that $\langle x \rangle = \langle p \rangle = 0$, we have $V_x + V_p = 2\langle \hat{n} \rangle + 1$. Solving for the larger variance, we find $N_F = 1/2 + \langle \hat{n} \rangle + \sqrt{\langle \hat{n} \rangle (\langle \hat{n} \rangle + 1)}$. For large $\langle \hat{n} \rangle$, this scales as $N_F \approx 2\langle \hat{n} \rangle$.

The CV mode analogue of the GHZ state is a superposition of coherent states, $|\psi\rangle \propto |\alpha\rangle + |-\alpha\rangle$, commonly referred to as a “cat state” [76]. Letting $\alpha \in \mathbb{R}$ without loss of generality, the largest quadrature variance is

$$V_x = \frac{1}{2} + \frac{2\alpha^2}{1 + e^{-2\alpha^2}}. \quad (5.21)$$

For large α , we have $N_F \approx 2\alpha^2 \approx \langle \hat{n} \rangle$.

For number states $|n\rangle$, the variance is the same for any quadrature, and gives $N_F = n + 1/2$.

It is worth briefly mentioning the role of the uncertainty relation here. For pure states, a small variance in one quadrature guarantees a large coherence in the conjugate quadrature – this is the case for squeezed states. In fact, this can be generalised to mixed states by a tighter form of the uncertainty relation [212]

$$\mathcal{F}(\rho, A)V(\rho, B) \geq \langle i[A, B] \rangle_\rho^2. \quad (5.22)$$

However, this is not always useful as a witness of large QFI; both cat states and number states demonstrate this. In general, $\mathcal{F}(\rho, x)$ measures the sensitivity of a state to translations along the conjugate direction p ; thus the QFI will be large whenever the p distribution has narrow features. This is the case for cat states and number states, without the need for a small variance in p .

5.5.2 Measure by Lee and Jeong

A measure was proposed in [70] to quantify macroscopicity in phase space for CV variables. Their motivation was rather similar to the intuition discussed at the end of the previous section: highly non-classical states generally have fast-oscillating features in their Wigner function. The measure adds up the sizes of each Fourier component in the Wigner function, giving a higher weight to larger frequencies. For N modes, we define a displacement operator $D(\boldsymbol{\alpha}) := \prod_i e^{\alpha_i a_i^\dagger - \alpha_i^* a_i}$ with $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \dots, \alpha_N)$. Then the measure is¹

$$N_{LJ}(\rho) := \frac{1}{2\pi^N} \int d^{2N} \boldsymbol{\alpha} |\boldsymbol{\alpha}|^2 |\chi_\rho(\boldsymbol{\alpha})|^2, \quad (5.23)$$

where $|\boldsymbol{\alpha}|^2 := \sum_{i=1}^N |\alpha_i|^2$ and $\chi_\rho(\boldsymbol{\alpha}) := \text{Tr}[\rho D(\boldsymbol{\alpha})]$. Note that χ_ρ is the Fourier transform of the Wigner function [76] (Chapter 3). The following results will be of use later:

Result 16. *The measure N_{LJ} has the following properties:*

¹The original definition used $|\boldsymbol{\alpha}|^2 - N$ in place of $|\boldsymbol{\alpha}|^2$. Removing the constant shift ensures that the measure is non-negative [213].

(i) It measures the rate of loss of purity under a decoherence model:

$$N_{LJ}(\rho) = -\frac{1}{2}\partial_t \text{Tr}(\rho^2), \quad (5.24)$$

$$\partial_t \rho = \mathcal{L}(\rho) := -\frac{1}{4} \sum_{i=1}^N [x_i, [x_i, \rho]] + [p_i, [p_i, \rho]]. \quad (5.25)$$

(A connection with a slightly different decoherence model was given in [70].)

(ii) Defining the QFI-like quantity $I_L(\rho, A) := -\frac{1}{2} \text{Tr}([A, \rho]^2)$, we can write

$$N_{LJ}(\rho) = \frac{1}{2} \sum_{i=1}^N [I_L(\rho, x_i) + I_L(\rho, p_i)]. \quad (5.26)$$

(iii) By averaging $I_L(\rho, x^\theta)$ over all angles θ , we have

$$N_{LJ}(\rho) = \sum_{i=1}^N \int \frac{d\theta}{2\pi} I_L(\rho, x_i^\theta). \quad (5.27)$$

Proof. Part (i) follows from first showing that $\partial_t \chi_\rho(\boldsymbol{\alpha}) = -\frac{1}{2} |\boldsymbol{\alpha}|^2 \chi_\rho(\boldsymbol{\alpha})$ under the given model. The remainder of the straightforward but lengthy calculation is given in Appendix B.

For (ii), we can use (a) to show $N_{LJ}(\rho) = -\text{Tr}(\rho \partial_t \rho) = -\text{Tr}[\rho \mathcal{L}(\rho)]$, and then use $\text{Tr}(\rho[A, [A, \rho]]) = \text{Tr}([A, \rho]^2)$.

Finally, for (iii), note that the results in (i) and (ii) are isotropic – x_i, p_i can be replaced with $x_i^{\theta_i}, p_i^{\theta_i}$ for any θ_i . Then we get the desired equation by averaging both sides of (5.26) over all angles. $\square$

The significance of parts (ii) and (iii) is that we have extracted a QFI-like quantity from the measure: $I_L(\rho, A)$ measures the size of the commutator $[A, \rho]$, just as do the generalised QFI quantities defined in Section 5.3.3. However, I_L is not a member of that set. In order to make N_{LJ} more similar in form to N_F , let us define a refined version

$$N'_{LJ}(\rho) := \max_{A \in \mathcal{A}} \frac{I_L(\rho, A)}{N}. \quad (5.28)$$

Here, the maximisation is over all quadratures – but this can be replaced by any other set of macroscopic observables, such as sums of Pauli operators for N qubits, in order to generalise to arbitrary systems in the same way as N_F . Given that the original quantity can be expressed as an average coherence over all directions, we have

$$N_{LJ}(\rho) \leq N N'_{LJ}(\rho). \quad (5.29)$$

The normalisation factor of N ensures that N'_{LJ} has the same $O(1)$ scaling behaviour as N_F for product states.

We now make a more in-depth comparison between I_L and the QFI.

Result 17. *Properties of I_L :*

- (i) *reduces to the variance for pure states:* $I_L(|\psi\rangle\langle\psi|, A) = V(|\psi\rangle\langle\psi|, A)$;
- (ii) *in general is a lower bound to the QFI:* $I_L(\rho, A) \leq \frac{1}{4}\mathcal{F}(\rho, A)$;
- (iii) *can be written in terms of the **Hilbert-Schmidt distance** $D_{HS}(\rho, \sigma) := \text{Tr}[(\rho - \sigma)^2]$ as*

$$I_L(\rho, A) = \frac{1}{4}\partial_t^2 D_{HS}(\rho, \mathcal{U}_t(\rho))|_{t=0}; \quad (5.30)$$

- (iv) *satisfies (M1), (M3) and (M4) but not (M2).*

Proof. For (i), we simply rearrange $I_L(\rho, A) = -\frac{1}{2}\text{Tr}(\rho A \rho A + A \rho A \rho - \rho A^2 \rho - A \rho^2) = \text{Tr}(\rho^2 A^2) - [\text{Tr}(\rho A)]^2$, and then use the fact that pure states satisfy $\rho^2 = \rho$.

For (ii), expanding $\rho = \sum_i \lambda_i |\psi_i\rangle\langle\psi_i|$ in its eigenbasis gives $I_L(\rho, A) = \frac{1}{2} \sum_{a,b} (\lambda_a - \lambda_b)^2 |\langle\psi_i|A|\psi_j\rangle|^2$. This can be compared against the corresponding expression for the QFI (5.7), using $(\lambda_i + \lambda_j)^{-1} \geq 1$.

To $O(t^2)$, we have $\mathcal{U}_t(\rho) - \rho \approx -it[A, \rho] + t^2 A \rho A - t^2/2 \{A^2, \rho\}$. Then $D_{HS}(\rho, \mathcal{U}_t(\rho)) = -t^2 \text{Tr}([A, \rho]^2) + O(t^3)$, from which (iii) follows.

Condition (M1) is clear since $I_L(\rho, A) = 0$ iff $[A, \rho] = 0$. (M4) is immediate from part (i). (M3) requires a calculation for $p\rho + q\sigma$, where $q = 1 - p$:

$$\begin{aligned} I_L(p\rho + q\sigma, A) &= p^2 \text{Tr}(\rho^2 A^2) - p^2 [\text{Tr}(\rho A)]^2 + q^2 \text{Tr}(\sigma^2 A^2) - q^2 [\text{Tr}(\sigma A)]^2 \\ &\quad + pq \text{Tr}(\{\rho, \sigma\} A) - 2pq \text{Tr}(\rho A) \text{Tr}(\sigma A) \\ &= pI_L(\rho, A) + qI_L(\sigma, A) - pq \text{Tr}([\rho^2 + \sigma^2 - \{\rho, \sigma\}] A^2) \\ &\quad - pq [\text{Tr}(\rho A) - \text{Tr}(\sigma A)]^2 \\ &= pI_L(\rho, A) + qI_L(\sigma, A) - pq I_L(\rho - \sigma, A) \\ &\leq pI_L(\rho, A) + qI_L(\sigma, A). \end{aligned} \quad (5.31)$$

The last line follows from observing that $I_L(X, A) = \frac{1}{2} \text{Tr}((-i[A, X])^2) \geq 0$ for any Hermitian operator X , as $-i[A, X]$ is also Hermitian.

To see that (M2) fails, suppose we start with the system S and an ancilla α in a state $\rho_S \otimes \sigma_\alpha$, taking the observable of interest to be trivial on α , i.e. just A_S . Then $I_L(\rho_S \otimes \sigma_\alpha, A_S) = I_L(\rho_S, A_S) \text{Tr}(\sigma_\alpha^2)$. As long as the ancilla state is not pure, $\text{Tr}(\sigma_\alpha^2) < 1$. Hence, by tracing out α – which is a free operation – I_L increases. $\square$

We have shown here that I_L is a faithful witness to the QFI. In fact, this lower bound property (for which the “L” stands) was used in [201] to provide a measurable witness for the Wigner-Yanase-Dyson skew information. While it equals the QFI for pure states, we have shown that it fails to be a monotone under the free operations. This stems from its relation to the Hilbert-Schmidt distance, which is known to be similarly problematic for defining other resource-theoretic measures, including entanglement and discord, because it is not contractive under quantum channels [182, 214].²

In the case of N qubits, we have seen that a scaling of $O(N^2)$ versus $O(N)$ scaling of the QFI can be considered a distinction between macroscopic and non-macroscopic quantumness. Since $I_L \leq \frac{1}{4}\mathcal{F}$, non-macroscopic scaling of $\mathcal{F}$ implies the same for I_L . Is the reverse true? If so, I_L could still be useful for distinguishing between these qualitative behaviours. In fact, we find this is false, giving a counterexample using a highly mixed state:

Result 18. *There exists a family of states on N qubits for which $I_L(\rho_N, Z) = O(N)$ but $\mathcal{F}(\rho_N, Z) = O(N^2)$.*

Proof. Let N be even, and define

$$\begin{aligned} \rho &:= \sum_{k=0}^{N/2-1} \frac{2}{N} |\psi_k^N\rangle \langle \psi_k^N|, \\ |\psi_k\rangle &:= \frac{1}{\sqrt{2}} (|0\rangle^{\otimes(N-k)} |1\rangle^{\otimes k} + |1\rangle^{\otimes(N-k)} |0\rangle^{\otimes k}). \end{aligned} \quad (5.32)$$

Note that the $|\psi_k\rangle$ are orthogonal; a complete basis can be made by including $|\phi_k\rangle := (|0\rangle^{\otimes(N-k)} |1\rangle^{\otimes k} - |1\rangle^{\otimes(N-k)} |0\rangle^{\otimes k})/\sqrt{2}$. This is also an eigenbasis for ρ . One can easily verify the following matrix elements:

$$\begin{aligned} \langle \psi_k | Z | \psi_l \rangle &= \langle \phi_k | Z | \phi_l \rangle = 0 \\ \langle \psi_k | Z | \phi_l \rangle &= \delta_{k,l} (N - 2k). \end{aligned} \quad (5.33)$$

²A rough fix could be attempted by dividing by the purity $\text{Tr}(\rho^2)$. However, this has been tried for N_{LJ} and leads to unsatisfactory results [213].

Using the expansion (5.7), we have

$$\begin{aligned}\frac{1}{4}\mathcal{F}(\rho, Z) &= \sum_{k=0}^{N/2-1} \frac{2}{N} (N-2k)^2 \\ &= \frac{1}{3}(N+1)(N+2) = O(N^2).\end{aligned}\tag{5.34}$$

On the other hand, the corresponding expression for I_L gives

$$\begin{aligned}I_L(\rho, Z) &= \sum_{k=0}^{N/2-1} \left(\frac{2}{N}\right)^2 (N-2k)^2 \\ &= \frac{2}{3N}(N+1)(N+2) = O(N).\end{aligned}\tag{5.35}$$

□

5.5.3 Measure by Nimmrichter and Hornberger

The measure presented in [54] is of a rather different character from those examined so far; however, we show here that there is a strong connection. The starting premise is that macroscopic quantum states should be viewed as tools for testing physics beyond standard quantum mechanics. Without assuming any particular physical mechanism, the authors argue from a few physical principles³ that the resulting modification to the Schrödinger equation should take a certain form. Without giving here the exact (rather complex) form, we state it for a point-like object of mass M moving in one dimension:

$$\partial_t \rho = -\frac{i}{\hbar} [H, \rho] + \mathcal{L}_1(\rho),\tag{5.36}$$

$$\mathcal{L}_1(\rho) = \frac{1}{\tau} \left[\int ds dq g(s, q) D(s, q) \rho D(s, q)^\dagger - \rho \right],\tag{5.37}$$

where H is the internal Hamiltonian of the system and $D(s, q) := e^{i(sp-qx)/\hbar}$ is the displacement operator.

The parameter τ , which is an effective coherence time for the object, and function $g(s, q)$, both depend on the mass M and structure of the object, as well as some free parameters. The shape of $g(s, q)$ is set by a fundamental reference distribution

³These are that the modification should (i) be covariant under Galilean transformations (in the non-relativistic regime), (ii) respect particle exchange symmetry, (iii) be unchanged when adding an uncorrelated system, and (iv) display scale-invariance (so that a rigid body can be treated as a point particle).

$g_e(s, q)$, which is used in the description for a single electron – it is chosen to be a Gaussian for simplicity. The standard deviations of s and q for $g(s, q)$ are $\frac{m_e}{M}\sigma_s$ and σ_q respectively, where m_e (the electron mass), σ_s and σ_q are some of the reference constants. In order to remain in the non-relativistic regime, the restrictions $\sigma_s \lesssim 20$ pm, $\hbar/\sigma_q \gtrsim 10$ fm are imposed. There is also a final free parameter τ_e which sets the coherence time scale for a single electron.

The modification $\mathcal{L}_1$ causes the state to “diffuse” in phase space, with a strength proportional to $1/\tau$ and a range set by $g(s, q)$. This also results in decoherence in x and p : when the internal dynamics can be neglected ($H = 0$) we have

$$\partial_t \chi_\rho(s, q) = -\frac{1}{\tau} [1 - \tilde{g}(q, s)] \chi_\rho(s, q), \quad (5.38)$$

where $\chi_\rho(s, q) := \text{Tr}[\rho D(s, q)]$ and $\tilde{g}(q, s) := \int ds' dq' g(s', q') e^{-i(qs' - sq')/\hbar}$ [215]. Since $\tilde{g}(q, s)$ is bounded, the decoherence rate saturates at some maximum [54]. Whenever we are far from saturation (i.e. well within the non-relativistic regime), we have

$$\partial_t \chi_\rho(s, q) \approx -\frac{1}{2\tau} \left[\left(\frac{\sigma_s m_e}{\hbar M} \right)^2 q^2 + \left(\frac{\sigma_q}{\hbar} \right)^2 s^2 \right] \chi_\rho(s, q). \quad (5.39)$$

Using the results in Appendix B, and the fact that $\frac{1}{\tau} \propto M^2$ [215], it follows that

$$\partial_t \rho \approx -\text{const.} \times \left((\sigma_s m_e)^2 [p, [p, \rho]] + (\sigma_q M)^2 [x, [x, \rho]] \right). \quad (5.40)$$

Note that this is of the same form (up to constants) as the decoherence model given in Result 16a, for a single subsystem.

The measure proposed by Nimmrichter and Hornberger is not a function of the state: it is defined for a particular *experiment*. Suppose an experiment determines that coherence can be maintained in a macroscopic system for some time. This will rule out certain values of the unknown parameter τ_e . Let $\tau_e^{\min}$ be the minimum value of τ_e which is consistent with the experiment (i.e. the strongest possible decoherence mechanism which has not been ruled out). Then the measure of macroscopicity is

$$N_{NH} := \log_{10} \left(\frac{\tau_e^{\min}}{1 \text{ s}} \right). \quad (5.41)$$

So experiments that rule out more values of τ_e are assigned a larger macroscopicity value.

We note that, instead of looking at the range of parameter values excluded by an experiment, one could focus on the decoherence rate of the state produced in the

experiment. This now gives a value which is a function of the state; qualitatively, the idea is the same, as a faster decaying state will be more useful for testing the possible parameter ranges. A simple quantity to look at is the rate of decay of the purity $\text{Tr}(\rho^2)$. Comparing with the results in the previous section, we have

$$-\partial_t \text{Tr}(\rho^2) \propto (\sigma_q M)^2 I_L(\rho, x) + (\sigma_s m_e)^2 I_L(\rho, p), \quad (5.42)$$

which is essentially the same as N_{LJ} for a single mode.

Let us also make a remark about the general approach of [54]. In order to make estimates for a number of past and proposed experiments, the authors had to estimate the corresponding coherence times under the collapse model. This can be done unambiguously for a two-branch superposition by taking the decay rate of some matrix element $\langle 0|\rho|1\rangle$. However, in general, it is not clear how to uniquely define such a decoherence rate. Indeed, we would argue that the programme of research into coherence measures is directed at identifying such quantities.

5.6 Conclusions

We started this chapter from the premise that quantum macroscopicity can be quantified in terms of coherence – defined with respect to a suitable choice of macroscopic observable. In order to arrive at a resource theory of coherence that respects not just the eigenbasis of the preferred observable, but also its eigenvalues, the free operations have to be more restricted than those studied in the previous two chapters. The resulting theory turned out to be the resource theory of $U(1)$ or $(\mathbb{R}, +)$ asymmetry, where the free operations are those commuting with transformations generated by the preferred observable.

Having examined some examples of monotones in this theory, we have seen that the QFI plays a special role as the asymptotic formation cost of many copies of a state from pure coherence units. Moreover, we have shown that in many physical cases (for harmonic and continuous spectra), it is possible to find a single pure state coherence unit from which all other states can be deterministically obtained. Questions about distillation and asymptotic cost are left open.

These results provide rigorous justification for a recently proposed measure of macroscopicity N_F based on the QFI, which quantifies the maximal coherence with

respect to any macroscopic observable in a system. We have applied the same analysis to another measure N_{LJ} , showing that it can be viewed similarly, using another function I_L in place of the QFI. However, this function is deficient as a measure of coherence, failing to be monotonic under the free operations. This calls into question the usefulness of the measure N_{LJ} . The measure N_{NH} is defined rather differently as a characteristic of an experiment, rather than a state. However, we have shown that the decoherence model upon which it is based provides a connection with N_{LJ} , which is the rate of loss of purity under this model.

Certain questions arising from this discussion merit further attention, and we turn to these in Chapter 7. Specifically, the reason for the normalisation factor of N in the definition of N_F for N subsystems is presently rather ad-hoc. In addition, we have not yet demonstrated that different systems can really be meaningfully compared with a measure such as N_F .

Chapter 6

Macroscopic distinguishability

In Chapter 5, we justified the QFI as a measure of macroscopicity via resource-theoretic notions of state transformations in the asymptotic limit of many copies. As this is rather abstract, one might prefer to have a measure which has an interpretation in terms of the kinds of operations which are actually feasible when dealing with large systems (after all, generating one macroscopic superposition is difficult enough, let alone arbitrarily many!). Such a measure has been defined recently [51, 112], however it applies only to equally weighted two-branch superpositions $(|\psi_0\rangle + |\psi_1\rangle)/\sqrt{2}$. As in our general framework, this makes reference to a preferred observable A – and the aim is to quantify the extent to which the branches are macroscopically distinguishable in a measurement of this observable. In a sense, the quantumness (purity) is assumed to be given, while the “size” of the superposition with respect to A is to be quantified.

One considers a noisy measurement of A , such that its distribution is blurred by the addition of Gaussian noise with variance Δ^2 . Let us denote by $f_\Delta^0(x)$ the probability distribution of such a measurement on $|\psi_0\rangle$, and similarly for $|\psi_1\rangle$. Suppose that one of these two states has been prepared with equal probability of $1/2$. The distinguishability of the distributions can be quantified by the maximal probability of correctly guessing, in a single shot, which of the two was prepared. This is given by

$$P_\Delta(\psi_0, \psi_1) = \frac{1}{2} \left(1 + \frac{1}{2} \int dx |f_\Delta^0(x) - f_\Delta^1(x)| \right). \quad (6.1)$$

The measure of macroscopicity is then defined to be the maximum tolerable noise Δ such that the two states are still distinguishable to within some threshold probability P_0 , i.e. $\max\{\Delta \mid P_\Delta(\psi_0, \psi_1) \geq P_0\}$.

The work [51] then goes on to assess the fragility of macroscopic quantum states

using this measure. Here, they refer to the fragility of entanglement between a microscopic system and a macroscopic system under decoherence induced by an interaction with an environment. Such a “micro-macro” entangled state occurs in the original description of Schrödinger’s cat, in which a microscopic system (a decaying atom) effectively acts as a control qubit for the state of a macroscopic system (a cat). Instead of just a superposition, we then have an entangled state of the form $(|0\rangle|\psi_0\rangle + |1\rangle|\psi_1\rangle)/\sqrt{2}$. Micro-macro entangled states are being pursued in experiments such as [216], which entangles a single photon to an atomic ensemble. Measuring entanglement is one way of verifying quantumness in these set-ups. The results in [112] show that, in two-branch versions of these states, the susceptibility of entanglement to decoherence is directly related to the noisy distinguishability of the branches.

In this chapter, we extend this measure to arbitrary states, retaining the core principle that macroscopic superpositions are those whose branches are easily distinguishable even with a highly noisy measurement. We define two measures along these lines, one being close to the original proposal, and the other based on the quantum correlations (namely, basis-dependent discord) established between a system and a measurement device. After proving that both measures satisfy the criteria of Chapter 5, we demonstrate connections to the QFI in the limit of weak measurements, and show implications for the fragility of micro-macro entangled states.

6.1 Measurement model

Before defining the measures of coherence and macroscopicity, we discuss a model of a noisy measurement. The measurement device M is treated as a CV quantum system with a “pointer observable” x representing the real-valued quantity which one reads off the device to determine the measurement result. There is a conjugate momentum p satisfying the canonical commutation relation $[x, p] = i$. We assume that M is initially uncorrelated with the system S , in a pure state $\rho_M = |\chi_\Delta\rangle\langle\chi_\Delta|$ such that the pointer distribution is a Gaussian of mean zero and standard deviation Δ :

$$\langle x|\chi_\Delta\rangle = \sqrt{g_\Delta(x)}, \quad g_\Delta(x) := \frac{1}{\sqrt{2\pi}\Delta} e^{-\frac{x^2}{2\Delta^2}}. \quad (6.2)$$

The system and measurement device are coupled by a unitary interaction

$$\rho_S \otimes \rho_M \rightarrow \rho'_{SM} = e^{-itA_{SPM}}(\rho_S \otimes \rho_M)e^{itA_{SPM}}, \quad (6.3)$$

which is a controlled shift of the pointer x by an amount ta_i conditioned on the state $|i\rangle$ of the system. In the following, we set the interaction time $t = 1$ since any other value may be absorbed into Δ (using $\Delta \rightarrow \Delta/t$) without affecting the physics. The same is true of any interaction strength parameter that might multiply the interacting Hamiltonian. Therefore Δ can be thought of equivalently as the measurement noise (imprecision), or the inverse of the interaction strength or time. In the limit $\Delta \rightarrow 0$, the measurement is perfectly precise, and as $\Delta \rightarrow \infty$ it becomes increasingly weak.

After tracing out the measurement device, the resulting operation on the system is a partial dephasing channel $\rho_S \rightarrow \Phi^\Delta(\rho_S) := \text{Tr}_M \rho'_{SM}$. For the results below, recall that we are defining coherence with the free operations as TI operations with respect to the transformations $\mathcal{U}_k(\rho) = e^{-ikA} \rho e^{ikA}$.

Lemma 4. *The partial dephasing channel Φ^Δ has the following properties:*

- (i) $\Phi^\Delta = \int dk f_\Delta(k) \mathcal{U}_k$, where $f_\Delta(k) = \sqrt{\frac{2}{\pi}} \Delta e^{-\Delta^2 k^2}$ is the momentum-space probability distribution for $|\chi_\Delta\rangle$;
- (ii) $[\mathcal{E}, \Phi^\Delta] = 0$ for any TI channel $\mathcal{E}$;
- (iii) $\Phi^\Delta(|i\rangle\langle j|) = e^{-\frac{(a_i - a_j)^2}{8\Delta^2}} |i\rangle\langle j|$;
- (iv) $\Phi^\alpha \circ \Phi^\beta = \Phi^\gamma$, where $\gamma^{-2} = \alpha^{-2} + \beta^{-2}$.

Proof. For (i), we perform the partial trace over M using momentum eigenstates $|k\rangle$:

$$\begin{aligned}
\Phi^\Delta(\rho_S) &= \int dk \langle k|_M e^{-iA_{sPM}} (\rho_S \otimes |\chi_\Delta\rangle\langle\chi_\Delta|) e^{iA_{sPM}} |k\rangle_M \\
&= \int dk |\langle p|\chi_\Delta\rangle|^2 e^{-ikA_S} \rho_S e^{ikA_S} \\
&= \int dk f_\Delta(k) \mathcal{U}_k(\rho_S).
\end{aligned} \tag{6.4}$$

(ii) then immediately follows, since $[\mathcal{E}, \mathcal{U}_k] = 0$ by definition. Instead tracing out M

with position eigenstates, we have

$$\begin{aligned}
\Phi^\Delta(|i\rangle\langle j|) &= \int dx \langle x|_M e^{-iA_{SPM}} (|i\rangle\langle j|_S \otimes |\chi_\Delta\rangle\langle\chi_\Delta|_M) e^{iA_{SPM}} |x\rangle_M \\
&= |i\rangle\langle j| \int dx \langle x| e^{-ia_i p} |\chi_\Delta\rangle \langle\chi_\Delta| e^{ia_j p} |x\rangle \\
&= |i\rangle\langle j| \int dx \langle x - a_i | \chi_\Delta \rangle \langle\chi_\Delta | x - a_j \rangle \\
&= |i\rangle\langle j| \int dx \frac{1}{\sqrt{2\pi}\Delta} e^{-\frac{1}{4\Delta^2}[(x-a_i)^2 + (x-a_j)^2]} \\
&= e^{-\frac{(a_i - a_j)^2}{8\Delta^2}} |i\rangle\langle j|, \tag{6.5}
\end{aligned}$$

showing (iii). Part (iv) follows from this expression. $\square$

Note that the infinite precision limit $\Delta \rightarrow 0$ gives $\Phi^0 = \Phi$, the fully dephasing channel used in earlier chapters.

Finally, suppose we set $t = \Delta^{-2}$ – note that this is *not* the same as the time parameter in the measurement model. It is easily verified (by taking the $\langle i| |j\rangle$ matrix element) that partial dephasing can be constructed using a master equation with t as time parameter:

$$\partial_t \rho_t = -\frac{1}{8} [A, [A, \rho_t]], \quad \rho_t = \Phi^{1/\sqrt{t}}(\rho). \tag{6.6}$$

6.2 Construction

In order to generalise the above measure, we must first decide how to quantify the mutual distinguishability of an arbitrary number of branches in a superposition, with arbitrary weighting. We work with the following proposal by Sekatski [6] (see Fig. 6.1):

Definition 19. Let $|\psi\rangle = \sum_i c_i |i\rangle$, then the **distinguishability under noise Δ** of the branches of $|\psi\rangle$ is defined in the following way. Suppose Alice prepares an ensemble of eigenstates $|i\rangle$ of A with probabilities $p_i = |c_i|^2$, and sends these to Bob, who performs a measurement with Gaussian noise Δ . Then

$$\text{dist}_\Delta(|\psi\rangle, A) := J(A : B), \tag{6.7}$$

where $J(A : B)$ is the (classical) mutual information shared between Alice and Bob.

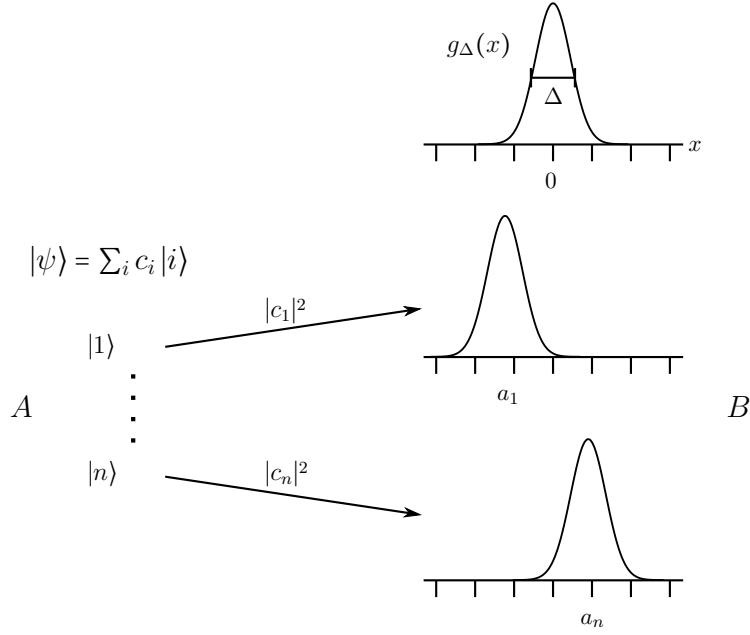


Figure 6.1: Setting for the definition of dist_Δ : given a pure state $|\psi\rangle = \sum_i c_i |i\rangle$, Alice sends the corresponding ensemble $|i\rangle$ with probabilities $|c_i|^2$ to Bob, who performs a measurement in order to distinguish them. His measurement consists of a Gaussian pointer of width Δ which shifts by a_i for each $|i\rangle$.

The motivation for this is the following: the more easily distinguishable the states are, the more information a noisy measurement gives about which branch was received. To be precise, the noisy measurement results in the joint statistics

$$p_{AB}(i, x) = p_i g_\Delta(x - a_i), \quad (6.8)$$

with marginals $p_A(i) = p_i$, $p_B(x) = \sum_i p_i g_\Delta(x - a_i)$. The mutual information is then easily shown to give

$$\begin{aligned} \text{dist}_\Delta(|\psi\rangle, A) &= H\left(\sum_i p_i g_\Delta(x - a_i)\right) - \sum_i p_i H(g_\Delta(x - a_i)) \\ &= H\left(\sum_i p_i g_\Delta(x - a_i)\right) - \frac{1}{2} \log(2\pi e \Delta^2), \end{aligned} \quad (6.9)$$

for which we have no simple analytical expression, but which is easy to calculate numerically. Some examples for N -qubit states are shown as a function of Δ in Fig. 6.2.

In order to extend the definition to mixed states, we use the convex roof construction presented in Chapter 5. This is a sensible approach because it guarantees that the measure vanishes for incoherent states.

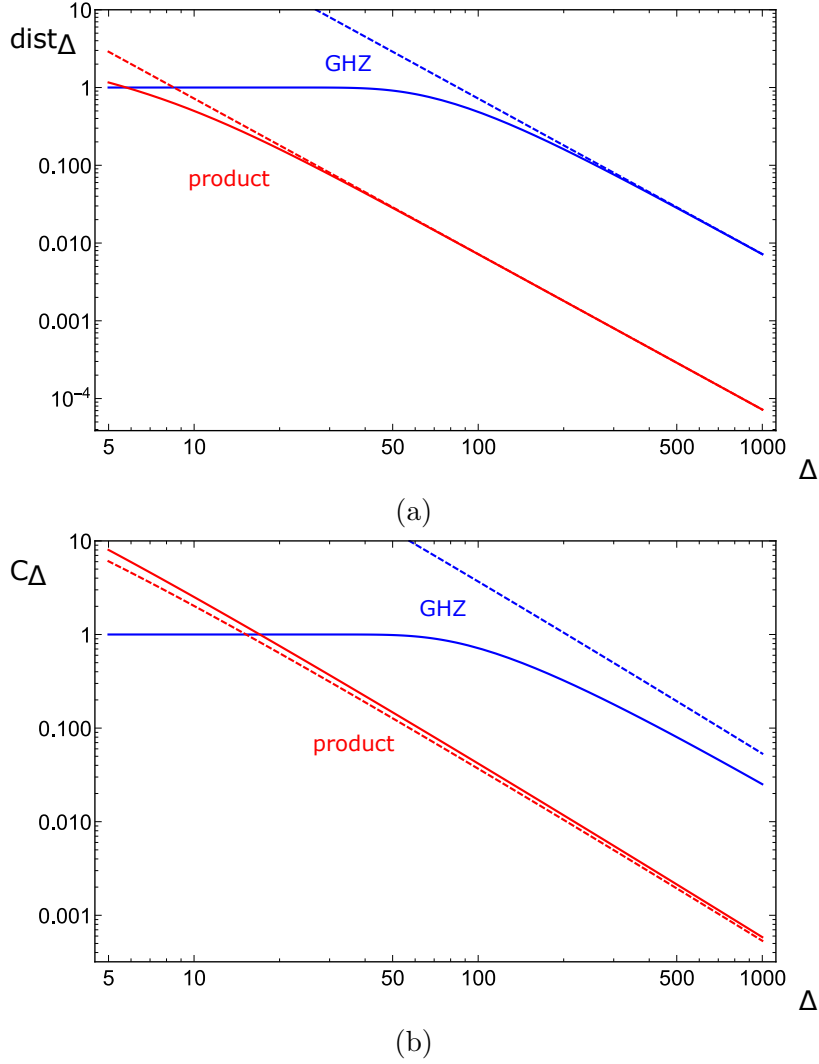


Figure 6.2: Plots of (a) $\text{dist}_\Delta(\rho, A)$ and (b) $C_\Delta(\rho, A)$ as a function of Δ , for $N = 100$ qubits with $A = \sum_{i=1}^N \sigma_i^z$ in a product state $|+\rangle^{\otimes N}$ (red solid line) and a GHZ state $(|0\rangle^{\otimes N} + |1\rangle^{\otimes N})/\sqrt{2}$ (blue solid line). Also shown (dashed lines) are the approximations (6.24) and (6.41) for dist_Δ and C_Δ respectively.

Definition 20. For an arbitrary state ρ , the distinguishability under noise Δ of the branches of ρ is defined by

$$\text{dist}_\Delta(\rho, A) := \min_{\{q_\alpha, |\psi_\alpha\rangle\}} \sum_\alpha q_\alpha \text{dist}_\Delta(|\psi_\alpha\rangle, A), \quad (6.10)$$

where the minimisation is over all pure state ensemble decompositions of ρ .

Now the macroscopicity measure is defined in the following way:

Definition 21. The *macroscopic branch distinguishability* of ρ is

$$\mu_\epsilon(\rho, A) := \min\{\Delta \mid \text{dist}_\Delta(\rho, A) \leq \epsilon\}, \quad (6.11)$$

where ϵ is a free parameter.

The idea behind this measure is that, as we increase the noise Δ , the branches become less distinguishable (this fact will be proved in the next section). At some point, the distinguishability drops to the threshold ϵ ; the more macroscopically distinct the branches, the greater the value of Δ for which this happens. Note that we could have instead used the definition $\max\{\Delta \mid \text{dist}_\Delta \geq \epsilon\}$, which is basically equivalent except that we would then have to prescribe $\mu_\epsilon = 0$ in the case that $\text{dist}_\Delta < \epsilon \forall \Delta$.

6.3 Properties

6.3.1 Coherence criteria

Here we show that the measures dist_Δ and μ_ϵ satisfy the necessary conditions to be good measures of macroscopic coherence. First, we give some useful properties of dist_Δ .

Lemma 5. $\text{dist}_\Delta(\rho, A)$ has the following properties:

- (i) strictly decreasing with respect to Δ (or else zero for all Δ);
- (ii) invariant under a change of scale $A \rightarrow \alpha A$, $\Delta \rightarrow \alpha \Delta$.

Proof. Note that both parts need only be shown for pure states, as the corresponding results for mixed states follow from the convex roof construction. For (i), rewrite (6.9) as $\text{dist}_\Delta(|\psi\rangle, A) = H(A + \Delta Z) - \frac{1}{2} \log(2\pi e \Delta^2)$, where Z is the standard normal distribution of mean zero and unit variance. Defining $\tau := \Delta^2$, we can apply **de Bruijn's identity** [86] (Theorem 17.7.2), which says

$$\partial_\tau H(A + \sqrt{\tau} Z) = \frac{1}{2 \ln 2} \mathcal{F}_c(A + \sqrt{\tau} Z), \quad (6.12)$$

where $\mathcal{F}_c$ is the classical translational Fisher information (2.41). Now we also use another result from statistics, derived by Stam [217]: for two random variables X and Y ,

$$\frac{1}{\mathcal{F}_c(X + Y)} \geq \frac{1}{\mathcal{F}_c(X)} + \frac{1}{\mathcal{F}_c(Y)}, \quad (6.13)$$

with equality iff both have a Gaussian distribution. Letting $X = A$, $Y = \sqrt{\tau} Z$, we see that $\mathcal{F}_c(A + \sqrt{\tau} Z) \leq \mathcal{F}_c(\sqrt{\tau} Z) = \frac{1}{\tau}$, where the final quantity is straightforwardly

computed from the integral (2.41). There is equality iff the distribution of A is the limiting case of a Gaussian with zero variance (so that $1/\mathcal{F}_c(A) = 0$).

So overall, we have $\partial_\tau[H(A + \sqrt{\tau}Z) - \frac{1}{2}\log(2\pi e\tau)] \leq \frac{1}{2\tau\ln 2} - \frac{1}{2\tau\ln 2} = 0$. When there is equality, A has zero variance, and so $H(A + \Delta Z) = \frac{1}{2}\log(2\pi e\Delta^2) \forall \Delta$.

For (ii), note that if the distribution of $A + \Delta Z$ is $f(x)$, then the corresponding distribution for $\alpha A + \alpha \Delta Z$ is $f(x/\alpha)/\alpha$. By a straightforward calculation, we have $H(\alpha A + \alpha \Delta Z) = H(A + \Delta Z) + \log \alpha$. The extra term cancels with the contribution from $\frac{1}{2}\log(2\pi e\alpha^2\Delta^2) = \frac{1}{2}\log(2\pi e\Delta^2) + \log \alpha$. $\square$

Result 19. $\text{dist}_\Delta(\rho, A)$ satisfies all conditions (M1-4).

Proof. We need only check (M1) for pure states, as a mixed state has $\text{dist}_\Delta(\rho, A) = 0$ iff there exists an ensemble decomposition with $\text{dist}_\Delta(|\psi_\alpha\rangle, A) = 0 \forall \alpha$. Now the concavity of the entropy tells us that $H(\sum_i p_i g_\Delta(x - a_i)) \geq \frac{1}{2}\log(2\pi e\Delta^2)$ with equality iff all the functions $g_\Delta(x - a_i)$ are the same, i.e. iff all a_i are the same for non-zero p_i . This is equivalent to saying that $|\psi\rangle$ is incoherent.

(M3) follows immediately from the convex roof definition.

(M2b) can first be shown for pure states. Suppose that a stochastic free operation takes $|\psi\rangle \rightarrow |\phi_\mu\rangle$ with probability w_μ , where $\sqrt{w_\mu}|\phi_\mu\rangle = K_\mu|\psi\rangle$, K_μ being a set of free Kraus operators of the form (3.34),

$$K_\mu = \sum_{i,j: a_i - a_j = \delta_\mu} c_{i,j}^\mu |i\rangle\langle j|. \quad (6.14)$$

Define $p_a := \langle \psi | \Pi_a | \psi \rangle$ and similarly $p_a^\mu := \langle \phi_\mu | \Pi_{a+\delta_\mu} | \phi_\mu \rangle$. Then it can be shown that $p_a = \sum_\mu w_\mu p_a^\mu$ (see Appendix C for the details); this is because each p_a^μ distribution is obtained by measuring and shifting p_a . The distribution of $A + \Delta Z$ for $|\psi\rangle$ is $f(x) = \sum_a p_a g_\Delta(x - a)$ – note that we have grouped together all terms i such that $a_i = a$. Similarly for $|\phi_\mu\rangle$ we have $f_\mu(x) = \sum_a p_a^\mu g_\Delta(x - a)$. Now $f(x) = \sum_\mu w_\mu f_\mu(x)$ and so the concavity of the entropy gives $H(f) \geq \sum_\mu w_\mu H(f_\mu)$. It follows that $\text{dist}_\Delta(|\psi\rangle, A) \geq \sum_\mu w_\mu \text{dist}_\Delta(|\phi_\mu\rangle, A)$, as required.

(M2a,b) are then immediately seen to hold as a consequence of convexity and (M2b) for pure states, just as argued for the QFI in Chapter 5.

Finally, for (M4), we let $|\psi\rangle = (|i\rangle + |j\rangle)/\sqrt{2}$. It is clear that $\text{dist}_\Delta(|\psi\rangle)$ is a function of only $|a_i - a_j|$ and Δ . So the requirement that $\text{dist}_\Delta(|\psi\rangle, A)$ be strictly increasing with $|a_i - a_j|$ is equivalent to it being strictly increasing under a replacement $A \rightarrow \alpha A$

with $\alpha > 1$. By part (ii) of Lemma 5, this change of scale can be transferred to $\Delta \rightarrow \Delta/\alpha$. So $\text{dist}_\Delta(|\psi\rangle, A)$ needs to be strictly decreasing with Δ – and this was shown in part (i). $\square$

So dist_Δ is a full convex monotone under TI operations, and respects the superposition scale criterion, making it suitable for measuring macroscopic coherence. While not all of these properties carry through to μ_ϵ , the most important ones do:

Result 20. $\mu_\epsilon(\rho, A)$ satisfies conditions (M2a) and (M4), as well as an ϵ -version of (M1):

$$\mu_\epsilon(\rho, A) = 0 \Leftrightarrow \text{dist}_\Delta(\rho, A) \leq \epsilon \ \forall \Delta. \quad (6.15)$$

Proof. For the vanishing criterion, note that either $\text{dist}_0(\rho, A) \leq \epsilon$ or else there exists a unique $\Delta > 0$ such that $\text{dist}_\Delta(\rho, A) = 0$ – this follows from it being strictly decreasing with Δ and from the plausible fact (see Result 21 below) that $\text{dist}_\Delta \rightarrow 0$ as $\Delta \rightarrow \infty$.

(M2a) for dist_Δ immediately implies the same for μ_ϵ . This is because $\text{dist}_\Delta(\mathcal{E}(\rho), A) \leq \text{dist}_\Delta(\rho, A)$ shows that the set of values of Δ for which $\text{dist}_\Delta(\mathcal{E}(\rho), A) \leq \epsilon$ is at least as big as the corresponding set for ρ .

Similarly, since $\text{dist}_\Delta(|\psi\rangle, A)$ is an increasing function of $|a_i - a_j|$ for $|\psi\rangle = (|i\rangle + |j\rangle)/\sqrt{2}$, the same is seen to be true for μ_ϵ . $\square$

6.3.2 Limiting cases

For $\Delta \rightarrow 0$, the measurement is infinitely precise and so the mutual information between Alice and Bob is

$$\text{dist}_0(|\psi\rangle, A) = J(A : B) = H(\{p_i\}) = S(\Phi(|\psi\rangle\langle\psi|)) = C_R(|\psi\rangle\langle\psi|). \quad (6.16)$$

Thus the convex roof construction gives

$$\text{dist}_0(\rho, A) = \min_{\{q_\alpha, |\psi_\alpha\rangle\}} \sum_\alpha q_\alpha C_R(|\psi_\alpha\rangle\langle\psi_\alpha|) = C_f(\rho), \quad (6.17)$$

which is the (speaking) coherence of formation. As shown in [117], this is equal to the asymptotic coherence cost. It satisfies $C_f(\rho) \geq C_R(\rho)$ (with equality for a set of states characterised in [117]). This limit evidently recovers a measure of speaking coherence.

We now turn to the opposite limit, where Δ is large. First we prove a statistical result which will be used later (this is in a sense dual to de Bruijn's identity; we have not seen it anywhere in the literature):

Lemma 6. *Let A, X be random variables and $B := X + tA$. For sufficiently small t ,*

$$H(B) = H(X) + \frac{t^2}{2\ln 2} \mathcal{F}_c(X) V(A) + O(t^3), \quad (6.18)$$

where $\mathcal{F}_c(X)$ is the classical translational Fisher information (2.41) of X , and $V(A)$ is the variance of A .

Proof. Denote the density functions of X, B by $g(x), f(x)$ respectively, and let A have values a_i with probabilities p_i (this is easily generalised to a continuous distribution). From the definition of B , we have

$$\begin{aligned} f(x) &= \sum_i p_i g(x - ta_i) \\ &\approx \sum_i p_i \left[g(x) - ta_i g'(x) + \frac{t^2 a_i^2}{2} g''(x) \right] \\ &= g(x) - t \langle A \rangle g'(x) + \frac{t^2 \langle A^2 \rangle}{2} g''(x), \end{aligned} \quad (6.19)$$

where we have done an expansion to $O(t^2)$. Similarly, it is easily shown that

$$\begin{aligned} f(x) \ln f(x) &\approx g(x) \ln g(x) + t [-\langle A \rangle g'(x) - \langle A \rangle g'(x) \ln g(x)] \\ &\quad + t^2 \left[\frac{\langle A \rangle^2 g'(x)^2}{2g(x)} + \frac{\langle A^2 \rangle}{2} g''(x) + \frac{\langle A^2 \rangle}{2} g''(x) \ln g(x) \right]. \end{aligned} \quad (6.20)$$

In order to find $H(Y)$, we have to use

$$\int_{-\infty}^{\infty} dx g'(x) (1 + \ln g(x)) = \left[g(x) (1 + \ln g(x)) \right]_{-\infty}^{\infty} - \int_{-\infty}^{\infty} dx g'(x) = 0, \quad (6.21)$$

$$\int_{-\infty}^{\infty} dx g''(x) (1 + \ln g(x)) = \left[g'(x) (1 + \ln g(x)) \right]_{-\infty}^{\infty} - \int_{-\infty}^{\infty} dx \frac{g'(x)^2}{g(x)} = -\mathcal{F}_c(X), \quad (6.22)$$

assuming g is sufficiently regular that $\lim_{x \rightarrow \pm\infty} g'(x) \ln g(x) = 0$. Hence we have

$$- \int_{-\infty}^{\infty} dx f(x) \ln f(x) \approx - \int_{-\infty}^{\infty} dx g(x) \ln g(x) + \frac{t^2}{2} \mathcal{F}_c(X) (\langle A^2 \rangle - \langle A \rangle^2), \quad (6.23)$$

from which the result follows. $\square$

Result 21. For all Δ , we have

$$\text{dist}_\Delta(\rho, A) \leq \frac{\mathcal{F}(\rho, A)}{(8 \ln 2) \Delta^2} \quad (6.24)$$

and in the limit of large Δ , these quantities become equal. Consequently,

$$\mu_\epsilon(\rho, A) \leq \sqrt{\frac{\mathcal{F}(\rho, A)}{(8 \ln 2) \epsilon}}, \quad (6.25)$$

with approximate equality for sufficiently small ϵ .

Proof. We calculate the relative entropy between the an arbitrary distribution $f(x)$ and the Gaussian $g_\Delta(x - \bar{x})$, where $\bar{x} := \langle x \rangle_f$:

$$\begin{aligned} H(f(x) \| g_\Delta(x - \bar{x})) &= -H(f) - \int dx f(x) \log g_\Delta(x - \bar{x}) \\ &= -H(f) + \frac{1}{2} \log(2\pi\Delta^2) + \int dx f(x) \frac{(x - \bar{x})^2}{(2 \ln 2) \Delta^2} \\ &= -H(f) + \frac{1}{2} \log(2\pi\Delta^2) + \frac{\langle (x - \bar{x})^2 \rangle_f}{(2 \ln 2) \Delta^2}. \end{aligned} \quad (6.26)$$

Choosing $f(x) = \sum_i |c_i|^2 g_\Delta(x - a_i)$ for a pure state $|\psi\rangle = \sum_i c_i |i\rangle$ gives

$$\begin{aligned} H(f(x) \| g_\Delta(x - \bar{x})) &= -H(f) + \frac{1}{2} \log(2\pi\Delta^2) + \frac{V(|\psi\rangle\langle\psi|, A) + \Delta^2}{(2 \ln 2) \Delta^2} \\ &= -\text{dist}_\Delta(|\psi\rangle, A) + \frac{V(|\psi\rangle\langle\psi|, A)}{(2 \ln 2) \Delta^2}, \end{aligned} \quad (6.27)$$

where we have used the fact that $f(x)$ is a convolution of p_i and $g_\Delta(x)$, under which the variance is additive. The first inequality (6.24) follows from the non-negativity of the relative entropy, and the fact that the QFI is the convex roof of the variance.

From Lemma 5, dist_Δ in the limit of large Δ is equivalent to taking small t in $H(tA + Z) - \frac{1}{2} \log(2\pi e/t^2)$, where $t = 1/\Delta$. Then Lemma 6 gives

$$\text{dist}_\Delta(|\psi\rangle, A) \approx \frac{t^2}{2 \ln 2} \mathcal{F}_c(Z) V(|\psi\rangle\langle\psi|, A) = \frac{t^2 V(|\psi\rangle\langle\psi|, A)}{(2 \ln 2) \Delta^2}. \quad (6.28)$$

Again, the convex roof gives the desired equation with the QFI.

The inequality (6.24) shows that if $\text{dist}_\Delta(\rho, A) = \epsilon$, then

$$\Delta^2 \leq \frac{\mathcal{F}(\rho, A)}{(8 \ln 2) \epsilon}, \quad (6.29)$$

giving the final inequality (6.25). Evidently, when ϵ is small, the required value of Δ is large, so the inequality is approximately saturated. $\square$

To summarise, this shows that the square root of the QFI provides an upper bound to the macroscopic branch distinguishability, and in the limit of a very weak (noisy) measurement, these quantities are proportional. Also note that the result of Lemma 6 only depends on $\mathcal{F}_c(X)$. So in the limit of a weak measurement, there is an equivalence between all pointer distributions of the same classical Fisher information. All of the above statements in this limit could be phrased in terms of an arbitrary (not just Gaussian) pointer, replacing Δ with $\mathcal{F}_c(X)^{-1/2}$ – by the classical Cramér-Rao bound, this quantity could be interpreted as the pointer sensitivity.

6.4 Alternative measure using discord

In this section, we build up an alternative measure with similar properties but a slightly different motivation. Recall that $\text{dist}_\Delta(|\psi\rangle, A)$ is defined as the *classical* mutual information between Alice, who prepares an ensemble of eigenstates of A , and Bob, who tries to distinguish them with a noisy measurement. A different but conceptually similar approach is to look at the correlations between the system S and measurement device M , using the model of Section 6.1. This could be equally well applied to initial pure and mixed states ρ_S . However, it would not distinguish between superpositions and incoherent mixtures; we must take only the quantum contribution to the correlations. A natural and simple way to do this is with the BDD as introduced in Chapter 4.

Definition 22. *The measure $C_\Delta(\rho, A)$ is defined as the BDD built up between system S and measurement device M in the model (6.3):*

$$C_\Delta(\rho, A) := \delta_b(M|S)_{\rho'} = I(S : M)_{\rho'} - J_b(M|S)_{\rho'}. \quad (6.30)$$

The following result shows that this actually has a simple expression as the von Neumann entropy generated by partial dephasing:

Result 22. $C_\Delta(\rho, A) = S(\Phi^\Delta(\rho)) - S(\rho)$.

Proof. Since the interaction is unitary, we have $S(\rho'_{SM}) = S(\rho_S \otimes \rho_M) = S(\rho_S) + S(\rho_M)$, thus $I(S : M)_{\rho'} = S(\rho'_S) + S(\rho'_M) - S(\rho_S) - S(\rho_M)$. For the classical correlations, we write $\Phi(\rho'_{SM}) = \sum_a p_a \rho_{S|a} \otimes \rho'_{M|a}$, where the sum is over distinct eigenvalues of A , and their corresponding projectors Π_a define $p_a \rho_{S|a} := \Pi_a \rho_S \Pi_a$ such that the $\rho_{S|a}$

are normalised. (Note that different $\rho_{S|a}$ are orthogonal.) We have also defined $\rho'_{M|a} := e^{-iap_M} \rho_M e^{iap_M}$, being the state of the measurement device corresponding to an eigenvalue a of A on the system. It follows that

$$\begin{aligned} J_b(M|S)_{\rho'} &= S(\Phi(\rho'_S)) + S(\rho'_M) - S(\Phi(\rho'_{SM})) \\ &= H(\{p_a\}) + \sum_a p_a S(\rho_{S|a}) + S(\rho'_M) - H(\{p_a\}) - \sum_a p_a [S(\rho_{S|a}) + S(\rho'_{M|a})] \\ &= S(\rho'_M) - \sum_a p_a S(\rho'_{M|a}). \end{aligned} \quad (6.31)$$

Since $\rho'_{M|a}$ is unitarily related to ρ_M , we have $J_b(M|S)_{\rho'} = S(\rho'_M) - S(\rho_M)$. Finally,

$$\delta_b(M|S)_{\rho'} = S(\rho'_S) - S(\rho_S) = S(\Phi^\Delta(\rho_S)) - S(\rho_S). \quad (6.32)$$

□

In addition, we have:

Lemma 7. C_Δ has the following properties:

- (i) $C_\Delta(\rho, A) \geq S(\rho \| \Phi^{\Delta/\sqrt{2}}(\rho))$;
- (ii) $C_\Delta(\rho, A) = \int dk f_\Delta(k) S(\mathcal{U}_k(\rho) \| \Phi^\Delta(\rho))$ (with $f_\Delta(k)$ as in Lemma 5);
- (iii) strictly decreasing with respect to Δ (or else zero for all Δ);
- (iv) invariant under a change of scale $A \rightarrow \alpha A$, $\Delta \rightarrow \alpha \Delta$.

Proof. (i) We need to use the fact that, for any quantum channel $\mathcal{N}$, $\text{Tr}[\mathcal{N}(\rho) \log \sigma] \leq \text{Tr}[\rho \log \mathcal{N}^\dagger(\sigma)]$, which is a consequence of the concavity of the logarithm [218] (Lemma 3.6). From this, we have

$$\begin{aligned} C_\Delta(\rho, A) &= -\text{Tr}[\Phi^\Delta(\rho) \log \Phi^\Delta(\rho)] - S(\rho) \\ &\geq -\text{Tr}[\rho \log \Phi^\Delta \circ \Phi^\Delta(\rho)] - S(\rho) \\ &= -\text{Tr}[\rho \log \Phi^{\Delta/\sqrt{2}}(\rho)] - S(\rho) \\ &= S(\rho \| \Phi^{\Delta/\sqrt{2}}(\rho)), \end{aligned} \quad (6.33)$$

having used Lemma 4 (iv) for the third line.

(ii) From Lemma 4(i),

$$\begin{aligned}
C_\Delta(\rho, A) &= -\text{Tr} \left[\int dk f_\Delta(k) \mathcal{U}_k(\rho) \log \Phi^\Delta(\rho) \right] - S(\rho) \\
&= \int dk f_\Delta(k) \left(-\text{Tr}[\mathcal{U}_k(\rho) \log \Phi^\Delta(\rho)] - S(\rho) \right) \\
&= \int dk f_\Delta(k) \left(-\text{Tr}[\mathcal{U}_k(\rho) \log \Phi^\Delta(\rho)] - S(\mathcal{U}_k(\rho)) \right) \\
&= \int dk f_\Delta(k) S(\mathcal{U}_k(\rho) \| \Phi^\Delta(\rho)), \tag{6.34}
\end{aligned}$$

where for the third line, we have used the fact that $\mathcal{U}_k$ leaves the entropy unchanged.

(iii) Given some $\Delta_1 < \Delta_2$, there exists $\alpha \in (0, \infty)$ such that $\Delta_1^{-2} = \Delta_2^{-2} + \alpha^{-2}$. Then, by Lemma 4 (iv), $\Phi^{\Delta_1} = \Phi^\alpha \circ \Phi^{\Delta_2}$. Therefore, taking d as the Hilbert space dimension,

$$S(\Phi^{\Delta_1}(\rho) \| I/d) = S(\Phi^\alpha \circ \Phi^{\Delta_2}(\rho) \| \Phi^\alpha(I/d)) \leq S(\Phi^{\Delta_2}(\rho) \| I/d), \tag{6.35}$$

using the monotonicity of the relative entropy. Since $S(\rho \| I/d) = \log d - S(\rho)$, this implies that $S(\Phi^{\Delta_1}(\rho)) \geq S(\Phi^{\Delta_2}(\rho))$. So C_Δ is decreasing with Δ .

Let us examine the case of equality in (6.35), applying the Petz recovery result (Theorem 1). The Petz recovery channel is simply $\mathcal{R} = \Phi^\alpha$ – this follows from using $\mathcal{E} = \mathcal{E}^\dagger = \Phi^\alpha$ and $\sigma = \mathcal{E}(\sigma) = I/d$ in Equation (4.48). Then the recovery condition becomes $\mathcal{R}[\Phi^\alpha \circ \Phi^{\Delta_2}(\rho)] = \Phi^{\Delta_2}(\rho)$, or equivalently $(\Phi^\alpha)^2 \circ \Phi^{\Delta_2}(\rho) = \Phi^{\Delta_2}(\rho)$. This is clearly satisfied iff $\Phi^{\Delta_2}(\rho) \in \mathcal{I}$. Therefore equality is attained iff $\rho \in \mathcal{I}$ (note that $\Delta_2 > 0$).

(iv) The scale-invariance is immediate from the fact that Φ^Δ multiplies the matrix element $|i\rangle\langle j|$ by a function of $(a_i - a_j)/\Delta$. $\square$

Result 23. C_Δ satisfies all criteria (M1-4).

Proof. If $\rho \in \mathcal{I}$, then $\Phi^\Delta(\rho) = \rho$, so $C_\Delta(\rho, A) = 0$. Conversely, Lemma 7 (i) shows that $C_\Delta(\rho, A) = 0 \Rightarrow S(\rho \| \Phi^{\Delta/\sqrt{2}}) = 0$ and so $\rho = \Phi^{\Delta/\sqrt{2}}(\rho)$. From Lemma 4 (iii), this must mean that all off-diagonal elements with differing a_i, a_j vanish – so $\rho \in \mathcal{I}$. Therefore (M1) holds.

Next, the expression in Lemma 7(ii) shows that, for any TI channel $\mathcal{E}$,

$$\begin{aligned}
C_\Delta(\mathcal{E}(\rho), A) &= \int dk f_\Delta(k) S(\mathcal{U}_k \circ \mathcal{E}(\rho) \| \Phi^\Delta \circ \mathcal{E}(\rho)) \\
&= \int dk f_\Delta(k) S(\mathcal{E} \circ \mathcal{U}_k(\rho) \| \mathcal{E} \circ \Phi^\Delta(\rho)) \\
&\leq \int dk f_\Delta(k) S(\mathcal{U}_k(\rho) \| \Phi^\Delta(\rho)) \\
&= C_\Delta(\rho, A),
\end{aligned} \tag{6.36}$$

having used $[\mathcal{E}, \Phi^\Delta] = 0$ and the monotonicity of the relative entropy. Therefore (M2a) holds; ensemble monotonicity (M2b) and convexity (M3) are also readily shown to be true in exactly the same way.

Finally, the argument for (M4) is exactly the same as with dist_Δ , using the scale-invariance property and the fact that C_Δ is decreasing with Δ . $\square$

Just as for dist_Δ , we can define a corresponding measure of macroscopicity by

$$\mu'_\epsilon(\rho, A) := \min\{\Delta \mid C_\Delta(\rho, A) \leq \epsilon\}. \tag{6.37}$$

Following exactly the same logic as before, we have:

Result 24. $\mu'_\epsilon(\rho, A)$ satisfies conditions (M2a), (M4) and an ϵ -version of (M1):

$$\mu'_\epsilon(\rho, A) = 0 \Leftrightarrow C_\Delta(\rho, A) \leq \epsilon \quad \forall \Delta. \tag{6.38}$$

In the strong measurement limit $\Delta \rightarrow 0$, we recover the relative entropy of coherence: $C_0(\rho, A) = S(\Phi(\rho)) - S(\rho) = C_R(\rho)$. We can also give a statement for the weak measurement limit which again gives a connection to the QFI, though not as directly as before.

Result 25. (i) Assume that ρ is full-rank (no zero eigenvalues). For sufficiently large Δ ,

$$C_\Delta(\rho, A) \approx \frac{\mathcal{F}_{KM}(\rho, A)}{4\Delta^2} \geq \frac{\mathcal{F}(\rho, A)}{(8 \ln 2)\Delta^2}, \tag{6.39}$$

where $\mathcal{F}_{KM}$ is the Kubo-Mori QFI [219] defined as

$$\mathcal{F}_{KM}(\rho, A) := \frac{1}{2} \sum_{i,j} (\lambda_i - \lambda_j) (\log \lambda_i - \log \lambda_j) |\langle \psi_i | A | \psi_j \rangle|^2. \tag{6.40}$$

(ii) If ρ is less than full-rank, then to leading order,

$$C_\Delta(\rho, A) \approx \frac{h(\Delta^{-2})}{4} \text{Tr}(\rho A^2 - P_\rho A \rho A), \tag{6.41}$$

where $h(t) := -t \log t$ and P_ρ is the projector onto the support of ρ . In particular, when ρ is pure, $\text{Tr}(\rho A^2 - P_\rho A \rho A) = V(\rho, A)$.

Proof. (i) Let $t := \Delta^{-2}$ be small, and $\sigma_t := \Phi^{1/\sqrt{t}}(\rho)$. Taking t as a time parameter, σ_t evolves according to the master equation (6.6). Under this model, there is a quantum equivalent of the de Bruijn identity for the rate of increase of the von Neumann entropy [220]:

$$\partial_t S(\sigma_t)|_{t=0} = -\text{Tr}[(\partial_t \log \rho) \log \sigma_t] = \frac{1}{4} \mathcal{F}_{KM}(\rho, A), \quad (6.42)$$

noting that $\sigma_0 = \rho$. From the expression (6.40), this derivative is finite only when all $\lambda_i > 0$. So for large Δ (small t), we have

$$C_\Delta(\rho, A) = S(\sigma_t) - S(\sigma_0) \approx \frac{t}{4} \mathcal{F}_{KM}(\rho, A). \quad (6.43)$$

This gives the claimed result. For the inequality, we compare the expansion (6.40) with the same for the usual QFI, (5.7). From an inequality between the so-called logarithmic mean and the arithmetic mean [221], we have, for any $x, y > 0$

$$\frac{x-y}{\ln x - \ln y} \leq \frac{x+y}{2} \Rightarrow \frac{1}{2}(x-y)(\log x - \log y) \geq \frac{(x-y)^2}{(\ln 2)(x+y)}, \quad (6.44)$$

hence $\mathcal{F}_{KM}(\rho, A) \geq \mathcal{F}(\rho, A)/(2 \ln 2)$.

(ii) We write the spectral decomposition $\rho = \sum_i \lambda_i |\psi_i\rangle\langle\psi_i|$, such that $\lambda_i > 0$ iff $i < r := \text{rank } \rho$. Let $\sigma_t = \sum_i \mu_i |\phi_i\rangle\langle\phi_i|$, with its eigenvalues and eigenstates implicitly functions of t and coinciding with those for ρ at $t = 0$. We also write their t -derivatives at $t = 0$ as $\dot{\mu}_i, |\dot{\phi}_i\rangle$. To lowest order, $\mu_i \approx \lambda_i + t\dot{\mu}_i$, so

$$\begin{aligned} S(\sigma_t) &= -\sum_i \mu_i \log \mu_i \\ &\approx -\sum_i (\lambda_i + t\dot{\mu}_i) \log(\lambda_i + t\dot{\mu}_i) \\ &= -\sum_{i < r} (\lambda_i + t\dot{\mu}_i) [\log \lambda_i + \log(1 + t\dot{\mu}_i/\lambda_i)] - \sum_{i \geq r} t\dot{\mu}_i [\log t + \log \dot{\mu}_i]. \end{aligned} \quad (6.45)$$

After the constant term $-\sum_i \lambda_i \log \lambda_i$, the leading order is $O(t \log t)$, so $S(\sigma_t) \approx S(\rho) - t \log t \sum_{i \geq r} \dot{\mu}_i$. Now,

$$\begin{aligned} \langle \psi_i | \partial_t \sigma_t | \psi_i \rangle &= \dot{\mu}_i + \lambda_i (\langle \psi_i | \dot{\phi}_i \rangle + \langle \dot{\phi}_i | \psi_i \rangle) \\ &= \dot{\mu}_i + \lambda_i \partial_t \langle \phi_i | \phi_i \rangle = \dot{\mu}_i. \end{aligned} \quad (6.46)$$

Since $\sum_i \mu_i = 1$ is constant, we have $\sum_{i \geq r} \dot{\mu}_i = -\sum_{i < r} \dot{\mu}_i = -\sum_{i < r} \langle \psi_i | \partial_t \sigma_t | \psi_i \rangle = -\text{Tr}(P_\rho \partial_t \sigma_t)$. Using the master equation (6.6), we find

$$\begin{aligned} \sum_{i \geq r} \dot{\mu}_i &= \frac{1}{8} \text{Tr} \left(P_\rho [A^2 \rho + \rho A^2 - 2A \rho A] \right) \\ &= \frac{1}{4} \text{Tr} (\rho A^2 - P_\rho A \rho A), \end{aligned} \quad (6.47)$$

having used the cyclic property of the trace and the fact that $P_\rho \rho = \rho P_\rho = \rho$. Putting these facts together gives the claimed result. For a pure state, $\rho = P_\rho = |\psi\rangle\langle\psi|$, so then $\text{Tr}(P_\rho A \rho A) = \langle \psi | A | \psi \rangle^2$. $\square$

See Fig. 6.2 for examples of C_Δ and the weak measurement approximation in the pure state case, where part (ii) applies.

6.5 Fragility of micro-macro entangled states

We focus on micro-macro entangled pure states of a microscopic system m with a macroscopic system S , of the form

$$|\Psi\rangle_{mS} = \sum_i \sqrt{p_i} |i\rangle_m |i\rangle_S, \quad (6.48)$$

where the $|i\rangle_m$ are arbitrary basis vectors for m and $|i\rangle_S$ are eigenstates of A_S , as usual. (Note that the coefficients can be assumed to be real and non-negative by absorbing phase factors into $|i\rangle_m$ – this form is then the Schmidt decomposition.) We examine the fragility of the entanglement under partial dephasing on the macroscopic part S , i.e.

$$\rho_{mS} = |\Psi\rangle\langle\Psi|_{mS} \rightarrow \rho'_{mS} = \Phi_S^\Delta(\rho_{mS}). \quad (6.49)$$

As noted earlier, we can also consider this dephasing to be generated by the master equation (6.6) with the time parameter $t = \Delta^{-2}$. The measurement device M now plays the role of an environment.

It is clear that such dephasing brings the state towards an S -incoherent state. In the hierarchy of quantum resources, vanishing coherence implies vanishing discord and entanglement; here, we provide quantitative statements to this effect.

The first result concerns the entanglement of formation E_f , which is defined as the convex roof of the entanglement entropy E_E :

Result 26. *Under the transformation (6.49),*

$$E_f(\rho'_{mS}) \leq E_f(\rho_{mS}) - \text{dist}_\Delta(\rho_{mS}, A_S). \quad (6.50)$$

This result was proved by Sekatski in [6] – we present the details in Appendix D for completeness. As a corollary, for any $\Delta \leq \mu_\epsilon(\rho_{mS}, A_S)$ – equivalently, $t \geq \mu_\epsilon(\rho_{mS}, A_S)^{-2}$ – the entanglement of formation is degraded by at least ϵ . In addition, using Result 21, for small t (large Δ),

$$E_f(\rho'_{mS}) \lesssim E_f(\rho_{mS}) - \frac{\mathcal{F}(\rho_{mS}, A_S)}{8 \ln 2} t. \quad (6.51)$$

Owing to the way we defined the measure C_Δ , it is plausible that it should let us say something about the degradation of coherence and of BDD between m and S . By a simple rearrangement, it is seen that

$$C_R^{m|S}(\rho_{mS}) - C_R^{m|S}(\rho'_{mS}) = C_\Delta(\rho_{mS}, A_S). \quad (6.52)$$

Statements about quantum correlations are also simple, since ρ_{mS} and ρ'_{mS} have structure of a **maximally correlated** state:

$$\rho_{AB} = \sum_{i,j} \rho_{ij} |i\rangle\langle j| \otimes |i\rangle\langle j|. \quad (6.53)$$

It has been noted [117, 118, 138] that there seems to be a correspondence between the resource theories of (speaking) coherence and of entanglement of maximally correlated states. Indeed, for a number of known instances, the entanglement of ρ_{AB} above equals a related measure of coherence of the state $\sum_{i,j} \rho_{ij} |i\rangle\langle j|$. This is certainly true for the relative entropies of entanglement and coherence – note that $E_f = E_d$, the distillable entanglement, for maximally correlated states [117]. Similarly, we can also see from (4.42), using the fact that the reduced states are incoherent, that the BDD of a maximally correlated state is simply $\delta_b(B|A)_\rho = C_R^{B|A}(\rho)$.

Partial dephasing of course leaves the diagonal part unchanged, so the change in both BDD and distillable entanglement between m and S is given by the change in von Neumann entropy:

$$\delta_b(m|S)_{\rho'} = \delta_b(m|S)_\rho - C_\Delta(\rho_{mS}, A_S), \quad (6.54)$$

$$E_d(\rho'_{mS}) = E_d(\rho_{mS}) - C_\Delta(\rho_{mS}, A_S). \quad (6.55)$$

Using Result 25 for small t , we also see that the fragility of E_d is directly related to the QFI of the initial state:

$$E_d(\rho'_{mS}) \approx E_d(\rho_{mS}) - \frac{\mathcal{F}(\rho_{mS}, A_S)}{4} h(t), \quad (6.56)$$

and similarly for $\delta_b(m|S)$.

6.6 Conclusions

In this chapter we have constructed measures of quantum macroscopicity from operational principles. This first aims to capture the mutual distinguishability of a set of branches of a superposition under a noisy measurement, and then is extended to mixed states in a way that takes into account coherent superpositions versus mixtures. The first quantity dist_Δ measures this distinguishability, using a Gaussian pointer distribution with uncertainty Δ . The second μ_ϵ then has an operational interpretation as the amount of noise Δ which renders the branches indistinguishable up to a threshold ϵ . We have shown that this construction fits well with the resource-theoretic axioms presented in Chapter 5, so the measures may be trusted as well-behaved unspeakable coherence quantifiers. In addition, a similar pair of measures C_Δ, μ'_ϵ have been derived by measuring the quantum correlations (BDD) built up between a system and measurement device.

By varying the free parameter Δ , we have seen that both dist_Δ and C_Δ interpolate between measures of speakable and unspeakable coherence. Strictly speaking, they are measures of speakable coherence only at $\Delta = 0$, satisfying condition (M4) for any $\Delta > 0$. For large Δ , corresponding to a very noisy measurement, dist_Δ is proportional to the QFI. This provides the QFI with a new meaning on top of its exposition in Chapter 5 as a formation cost. The same interpretation carries over to μ_ϵ for small ϵ . A similar but more complex result was found for C_Δ and μ'_ϵ – for full rank states, a member of the set of generalised QFI quantities is obtained, with the variance for pure states and a more intricate quantity for other cases.

We should note here that there are close parallels with a measure recently proposed in [222], based on the disturbance to a state caused by a coarse-grained measurement, modelled in the same way as we have done with a Gaussian pointer. Their measure is of the form $D(\rho, \Phi^\Delta(\rho))$, where D is any contractive pseudo-distance. This is

obviously similar to C_Δ in that the latter measures the amount of von Neumann entropy generated by the channel Φ^Δ . They also give a similar connection to the QFI, in the form of a bound $D_B(\rho, \Phi^\Delta(\rho)) \leq 2(1 - \exp[-I_{sk}(\rho, A)/(4\Delta^2)])$. Also noting (albeit in different language) that small Δ brings the measure closer to speakable coherence, they argue that this provides a case for showing that the criteria (M1-4) are not *sufficient* for ensuring a good measure of macroscopicity.

Finally, we have demonstrated the implications for the fragility of micro-macro entangled states under dephasing in the form of Φ^Δ , considered as a measurement performed by an environment. The quantities $\mu_\epsilon^{-2}, \mu'_\epsilon^{-2}$ can be interpreted as the tolerable coupling strength with the environment before an amount ϵ of entanglement is lost (as measured by E_f and E_d respectively). Using the earlier results about the large Δ limit, this also shows that the QFI is directly related to the fragility of this entanglement under dephasing. One topic for further work is to see if this can be generalised to more decoherence models. It would be instructive to characterise what kind of noise is most destructive, as this would guide our understanding of how to protect micro-macro entangled states from decoherence. Furthermore, in multipartite systems, one often encounters uncorrelated noise on each subsystem, in contrast to the model here where the environment couples globally to the additive observable A .

Chapter 7

Macroscopic coherence and entanglement of mechanical systems

The quantum Fisher information was identified in Chapter 5 as a significant quantity in discussing the macroscopic-scale quantum coherence of a system. However, the analysis has so far been rather generic in that we have not needed to specify any particular physical system. In the present chapter, we study some real experimental systems, aiming to show how the QFI can provide information about their macroscopic quantum behaviour. We specialise to the class of mechanical systems, which is broad enough to cover a range of recent relevant experiments, but also sufficiently narrow that meaningful comparisons can be expected to be drawn between different experiments. This may be seen as complementary to recent work studying photonic systems, trapped ions and atomic ensembles [223].

We argue that a full description of macroscopicity cannot be captured in a single number – as is often attempted in the literature. Instead, we define two independent quantities, both based on the QFI: extensive size and entangled size. The former is a measure of total coherence in atomic-scale units, while the latter counts the effective number of entangled subsystems, and is a witness of genuine many-body entanglement.

We estimate these quantities in systems including a thought experiment by Leggett, large diffracting molecules and micro-mechanical oscillators. While the numbers we obtain are order-of-magnitude estimates only, we provide a detailed example in the case of diffracting molecules to show how the calculations can be performed directly

from data obtained in the experiments. We compare the values obtained in different systems and offer some perspectives on how large macroscopicity values could be attained in the future.

7.1 Two elements of macroscopicity

7.1.1 General considerations

Given that we are focusing on mechanical systems, the preferred observables are chosen to be position and momentum. In particular, massive mechanical objects are most naturally described in terms of their total momentum and centre of mass position (as well as other possibilities such as angular momentum or the relative positions and momenta of macroscopic subsystems, which we neglect here). However, recall that our framework is based on additive observables – this holds for the total momentum P but not the centre of mass X_{CM} . To obtain an additive position-space variable, we multiply X_{CM} by the total mass. So for N subsystems of individual masses, positions and momenta m_i, x_i, p_i , we have the following preferred observables:

$$Q := \sum_{i=1}^N q_i = \sum_{i=1}^N m_i x_i = M X_{CM}, \quad (7.1)$$

$$P := \sum_{i=1}^N p_i \quad (7.2)$$

where $M = \sum_{i=1}^N m_i$ is the total mass. In principle, one could also take rotated quadratures as linear combinations of Q and P , but these two observables will be sufficient here. Note that, while $[X_{CM}, P] = i\hbar$, we have $[Q, P] = iM\hbar$.

To begin to identify the different types of macroscopicity quantity, we outline the main ideas introduced by Leggett [55]. For him, the first important aspect of a macroscopic superposition is its *extensive difference* – the difference in expectation value of some additive variable between two branches, expressed in some typical atomic-scale units. The first quantity defined below, the extensive size, encapsulates this idea in the framework of coherence. Moreover, it has the advantages of being defined for any state (not just a two-branch superposition) and taking into account the purity of the superposition in a quantitative way.

However, there are superpositions with a large extensive difference which one might be unwilling to call macroscopically quantum. Leggett gives the example of a

single neutron passing through an interferometer with branches separated by 10 cm. In units of Å, this has an extensive difference in position of around 10^9 . However, only a single particle is involved. He therefore suggests that we look at another quantity called the *disconnectivity*. Roughly speaking, this should count the number of entangled particles. The precise definition given in [65] is, in the words of its author, “from the point of view of modern quantum information theory...no doubt inelegant and could almost certainly be substantially improved” [55]. We propose an attempt at such an improvement with the entangled size defined below.

7.1.2 Extensive size

The first measure is a direct consequence of the discussion in Chapter 5 of the QFI as a macroscopic coherence preparation cost. Recalling Result 15 and Equation (5.16), the preparation of many copies of ρ from copies of a suitable coherence unit $|\psi_0\rangle$ costs $\mathcal{F}(\rho, A)/\mathcal{F}(|\psi_0\rangle\langle\psi_0|, A)$ per copy of ρ (via a prepare-and-mix strategy). Let us take $|\psi_0\rangle$ to have atomic-scale coherence in both position and momentum, so that $V(|\psi_0\rangle\langle\psi_0|, Q) = Q_0^2$, $V(|\psi_0\rangle\langle\psi_0|, P) = P_0^2$, with

$$Q_0 := m_u a_0, \quad (7.3)$$

$$P_0 := \frac{m_u \hbar}{Q_0} = \frac{\hbar}{a_0}, \quad (7.4)$$

where $m_u \approx 1.66 \times 10^{-27}$ kg is the unified atomic mass unit and $a_0 \approx 5.29 \times 10^{-11}$ m is the Bohr radius. Rather than maximising over the set of macroscopic observables, we will leave the choice of $A = Q$ or P open, and define:

Definition 23. The *extensive size* of a state ρ with respect to A is

$$N_{\text{ext}}(\rho, A) := \frac{\mathcal{F}(\rho, A)}{4A_0^2}. \quad (7.5)$$

To see the connection with Leggett’s extensive difference, consider a system of mass M in an equally weighted superposition $(|\psi_1\rangle + |\psi_2\rangle)/\sqrt{2}$, where the two branches are localised at positions $\pm\Delta x$. We assume that in each branch, the system is localised with uncertainty much less than Δx . Then we have

$$N_{\text{ext}}(\rho, Q) = \left(\frac{M\Delta x}{m_u a_0} \right)^2. \quad (7.6)$$

This is proportional to the *square* of the extensive difference Δx . This might seem surprising or undesirable at first, but follows from the fact that N_{ext} measures the total coherence of a state – and is thus additive with respect to independent copies. The variance (as opposed to the standard deviation) is also additive in this sense, so the total coherence must scale with $(\Delta x)^2$ rather than Δx .

It is worth noting that the decoherence rate in the model of Nimmrichter and Hornberger is closely related to the extensive size, from Equation (5.42):

$$\begin{aligned} -\partial_t \text{Tr}(\rho^2) &\propto (2\sigma_q M)^2 I_L(\rho, X_{CM}) + (2\sigma_s m_e)^2 I_L(\rho, P) \\ &\leq (\sigma_q Q_0)^2 N_{\text{ext}}(\rho, Q) + (\sigma_s m_e P_0)^2 N_{\text{ext}}(\rho, P). \end{aligned} \quad (7.7)$$

7.1.3 Entangled size

Now we turn to defining a measure which captures the effective number of entangled particles. In contrast to N_{ext} , we of course must first choose a partition of the system into subsystems. Such a choice is not unique, and in fact can have a substantial effect on the value of the measure – this was observed for the N_F measure in [69]. To motivate the definition, let us first examine the N -qubit measure by Fröwis and Dür (5.17), $N_F(\rho) = \max_{A \in \mathcal{A}} \mathcal{F}(\rho, A)/(4N)$. As remarked earlier, the normalisation factor of N ensures that $N_F \leq N$, attaining the maximum value for a full N -particle GHZ state. In fact, one major motivation for defining N_F is that it acts as a witness of multipartite entanglement. In order to state this, we must introduce the concept of a state in which no more than k parts are entangled [224]:

Definition 24. *A pure state $|\psi\rangle$ is called **k -producible** iff it can be written in the form*

$$|\psi\rangle = \bigotimes_{\alpha=1}^m |\psi_\alpha\rangle \quad (7.8)$$

such that each term $|\psi_\alpha\rangle$ is the state of a block of subsystems of size $k_\alpha \leq k$. A mixed state is then called k -producible iff it can be written as a mixture of k -producible pure states.

A state which is not k -producible is then considered to have genuine $(k+1)$ -partite entanglement. It was shown in [225, 226] that every k -producible state of N qubits satisfies $N_F(\rho) \leq k$. Thus, finding $N_F(\rho) > k$ is a witness of genuine $(k+1)$ -partite entanglement. However, note that this really is only a witness: $N_F \leq k$ does

not rule out genuine $(k + 1)$ -partite entanglement. In the pure state case, one can see that large QFI is a signature that all the A_i are positively correlated (this kind of entanglement is sometimes called “metrologically useful entanglement” in that it enables high precision metrology via the Cramér-Rao bound [227, 228]).

It is a non-trivial problem to extend this measure to arbitrary systems. For qubits, each A_i is a Pauli matrix and so has eigenvalues ± 1 . If we multiply A by a constant c , it is clear that the measure should correct for this rescaling by dividing by c^2 . So the denominator should include something along the lines of the squared spectral size of A_i . However, this is undefined in the case of CV modes, for which the spectrum is infinite. A reasonable guess is to use the local variance $V(\rho_i, A_i)$ (where $\rho_i = \text{Tr}_{N \setminus i} \rho$) as a measure of the (squared) spectral size *explored in the state* ρ . Therefore we define:

Definition 25. *The **entangled size** of a state ρ with respect to A is*

$$N_{\text{ent}}(\rho, A) := \frac{\mathcal{F}(\rho, A)}{4 \sum_{i=1}^N V(\rho_i, A_i)}. \quad (7.9)$$

We can also give this quantity the following resource-theoretic interpretation. Define a coherence unit $|\psi_0\rangle := \bigotimes_{i=1}^N |\psi_0^i\rangle$, such that each $V(|\psi_0^i\rangle\langle\psi_0^i|, A_i) = V(\rho_i, A_i)$; then $N_{\text{ent}}(\rho, A)$ is the asymptotic formation cost per copy of ρ from $|\psi_0\rangle$. In words, the reference state $|\psi_0\rangle$ is chosen to have the same local physical resources as ρ , in terms of the covered spectrum of A_i – but without any entanglement.

Moreover, N_{ent} is indeed a witness of genuine multipartite entanglement¹:

Result 27. *Any k -producible state ρ satisfies $N_{\text{ent}}(\rho, A) \leq k$.*

We break down the proof of this statement into a number of stages.

Lemma 8. (Maximum size) *Let ρ_α be a state of a block α of $k_\alpha \leq N$ subsystems, and define A_α to be the restriction of A to α . Then*

$$N_{\text{ent}}(\rho_\alpha, A_\alpha) \leq k_\alpha. \quad (7.10)$$

Proof. By the variance upper-bound to the QFI, we have

$$N_{\text{ent}}(\rho_\alpha, A_\alpha) \leq \frac{V(\rho_\alpha, A_\alpha)}{\sum_{i \in \alpha} V(\rho, A_i)}. \quad (7.11)$$

¹During preparation of this work, we became aware of a recent preprint [229] which defines the same quantity under the name of a “variance-assisted Fisher density”. However, they do not show that this is a multipartite entanglement witness, only that it detects entanglement.

It is simple to show that

$$V(\rho_\alpha, A_\alpha) = \sum_{i,j \in \alpha} \text{Cov}(\rho, A_i, A_j), \quad (7.12)$$

in terms of the covariance defined by $\text{Cov}(\rho, A_i, A_j) := \text{Tr}(\rho A_i A_j) - \text{Tr}(\rho, A_i) \text{Tr}(\rho, A_j)$ (bearing in mind that $[A_i, A_j] = 0$). The covariance satisfies the inequality

$$\text{Cov}(\rho, A_i, A_j) \leq \sqrt{V(\rho, A_i) V(\rho, A_j)}. \quad (7.13)$$

From this,

$$\begin{aligned} V(\rho_\alpha, A_\alpha) &\leq \sum_{i,j \in \alpha} \sqrt{V(\rho, A_i) V(\rho, A_j)} \\ &= \left(\sum_{i \in \alpha} \sqrt{V(\rho, A_i)} \right)^2 \\ &= k_\alpha^2 \left(\sum_{i \in \alpha} \frac{1}{k_\alpha} \sqrt{V(\rho, A_i)} \right)^2 \\ &\leq k_\alpha^2 \sum_{i \in \alpha} \frac{1}{k_\alpha} V(\rho, A_i) \\ &= k_\alpha \sum_{i \in \alpha} V(\rho, A_i), \end{aligned} \quad (7.14)$$

where the penultimate line uses the convexity of the square function. The result follows from putting this into (7.11). $\square$

Lemma 9. (*Tensor products*) *Let α, β be disjoint blocks of subsystems. Then*

$$N_{\text{ent}}(\rho_\alpha \otimes \rho_\beta, A_\alpha + A_\beta) \leq \max\{N_{\text{ent}}(\rho_\alpha, A_\alpha), N_{\text{ent}}(\rho_\beta, A_\beta)\}. \quad (7.15)$$

Proof. By the additivity of the QFI and variance,

$$\begin{aligned} N_{\text{ent}}(\rho_\alpha \otimes \rho_\beta, A_\alpha + A_\beta) &= \frac{\frac{1}{4} \mathcal{F}(\rho_\alpha, A_\alpha) + \frac{1}{4} \mathcal{F}(\rho_\beta, A_\beta)}{\sum_{i \in \alpha} V(\rho, A_i) + \sum_{j \in \beta} V(\rho, A_j)} \\ &= \frac{[\sum_{i \in \alpha} V(\rho, A_i)] N_{\text{ent}}(\rho_\alpha, A_\alpha) + [\sum_{j \in \beta} V(\rho, A_j)] N_{\text{ent}}(\rho_\beta, A_\beta)}{[\sum_{i \in \alpha} V(\rho, A_i)] + [\sum_{j \in \beta} V(\rho, A_j)]}. \end{aligned} \quad (7.16)$$

This takes the form of a weighted average of $N_{\text{ent}}(\rho_\alpha, A_\alpha)$ and $N_{\text{ent}}(\rho_\beta, A_\beta)$, which cannot exceed the larger of the two quantities. $\square$

Lemma 10. (*Mixing*) *Let ρ, σ be any states of the same system, and $p \in [0, 1]$. Then*

$$N_{\text{ent}}(p\rho + (1-p)\sigma, A) \leq \max\{N_{\text{ent}}(\rho, A), N_{\text{ent}}(\sigma, A)\}. \quad (7.17)$$

Proof. Simultaneously using the convexity of the QFI and the concavity of the variance [195], we have

$$\begin{aligned} N_{\text{ent}}(p\rho + (1-p)\sigma, A) &\leq \frac{\frac{1}{4}p\mathcal{F}(\rho, A) + \frac{1}{4}(1-p)\mathcal{F}(\sigma, A)}{\sum_i pV(\rho, A_i) + (1-p)V(\sigma, A_i)} \\ &= \frac{p[\sum_i V(\rho, A_i)]N_{\text{ent}}(\rho, A_i) + (1-p)[\sum_i V(\sigma, A_i)]N_{\text{ent}}(\sigma, A)}{p[\sum_i V(\rho, A_i)] + (1-p)[\sum_i V(\sigma, A_i)]}. \end{aligned} \quad (7.18)$$

This is again a weighted average of the individual effective sizes, so the result follows as above. $\square$

Now we are able to prove Result 27. From Lemmas 8 and 9, a k -producible pure state $|\psi\rangle = \bigotimes_{\alpha=1}^m |\psi_\alpha\rangle$ satisfies

$$N_{\text{ent}}(|\psi\rangle\langle\psi|, A) \leq \max\{k_\alpha \mid \alpha = 1, \dots, m\} \leq k. \quad (7.19)$$

Lemma 10 then gives the result for mixed k -producible states.

For qubits, N_{ent} can be a much more sensitive multipartite entanglement witness than N_F . For example, take the unbalanced GHZ state $|\psi\rangle = \sqrt{p}|0\rangle^{\otimes N} + \sqrt{1-p}|1\rangle^{\otimes N}$ and $A_i = \sigma_i^z$. Then we have

$$N_F(|\psi\rangle\langle\psi|, A) = 4p(1-p)N, \quad N_{\text{ent}}(|\psi\rangle\langle\psi|, A) = N. \quad (7.20)$$

So N_{ent} is able to detect genuine N -partite entanglement even for arbitrarily small p where N_F may fail. Incidentally, this shows that N_{ent} is not a continuous function of the state – in the limit $p \rightarrow 0$, $|\psi\rangle$ approaches a product state but N_{ent} has a constant value of N . Points of discontinuity are where the denominator vanishes (i.e., all $V(\rho_i, A_i) = 0$), since both the variance and QFI are continuous [230]. This is a set of vanishing volume, therefore we need not be concerned about the physical robustness of the measure. Also note that the mixing inequality of Lemma 10 is weaker than convexity.

7.2 Estimates for physical systems

7.2.1 Leggett's crystal

First we study a thought experiment in a recent note by Leggett [231] whose purpose was to demonstrate a deficiency in the measure of macroscopicity by Korsbakken et

al. [232] which was applied to superconducting qubits. Here, we use it as a toy model which will be useful in discussing the important concepts in real experiments below. Leggett wishes to define a state of matter which is just on the limit of what one would be intuitively comfortable to class as a macroscopic superposition. This is a block of solid matter – a crystal of LiF for the sake of concreteness – of side length $5\text{ }\mu\text{m}$, about the smallest size resolvable by the human eye. It is put into a superposition state $(|\psi_0\rangle + |\psi_1\rangle)/\sqrt{2}$ such that one branch is at rest while the other is moving at a speed $v = 5\text{ }\mu\text{ms}^{-1}$. The branches are assumed to have the same spatial distribution (although they of course separate over time). Our aim here is to demonstrate that intuition is in fact highly misleading in this case, and the state is far from resembling a cat state.

For simplicity, we model the crystal as being composed of atoms of a single mass $m \approx 12.5m_u$. We also ignore electrons, as the far greater mass of the nuclei means that they have the dominant contribution to both Q and P . The number of atoms is $N \approx 1.6 \times 10^{13}$. The extensive size in P is then $N_{\text{ext}}(\rho, P) = V(\rho, P)/P_0^2 = (\Delta P/P_0)^2 \approx 6.9 \times 10^{21}$, where $\Delta P = Mv/2 = Nmv/2 \approx 8.3 \times 10^{-19}\text{ kgms}^{-1}$.

However, the entangled size tells a different story. We take single atoms as the subsystems. Each atom is confined in the crystal such that its position uncertainty is of the order $\delta x \approx 10^{-11}\text{ m}$ (see [233] for typical values in solids at room temperature). The momentum uncertainty in a single branch is then expected to be $\delta p \approx \hbar/\delta x \approx 1.1 \times 10^{-23}\text{ kgms}^{-1}$. Including the uncertainty as a result of the superposition, we have $V(\rho_1, P) \approx (\delta p)^2 + (\Delta p)^2$, where $\Delta p = \Delta P/N \approx 5.2 \times 10^{-32}\text{ kgms}^{-1}$. Therefore

$$N_{\text{ent}}(\rho, P) \approx \frac{(\Delta P)^2}{N[(\delta p)^2 + (\Delta p)^2]} = N \left(\frac{r^2}{1 + r^2} \right), \quad r = \frac{\Delta p}{\delta p}. \quad (7.21)$$

The entangled size – and thus the degree to which the superposition resembles a cat state – depends crucially on the ratio r of the momentum change per atom between branches, to the momentum uncertainty per atom in a single branch. When r is large, each single-atom state is highly distinguishable between branches and so the state is well described as an N -particle cat state. At the other extreme – as is the case here, with $r \approx 5 \times 10^{-9}$ – each atom barely changes state between branches.² For the present example, $N_{\text{ent}}(\rho, P) \approx Nr^2 \approx 4 \times 10^{-4}$, which is negligible.

²This could be compared with a similar result about variants of N -qubit of GHZ states $|\psi_0\rangle^{\otimes N} +$

Due to the complete spatial overlap of the branches, both $N_{\text{ext}}(\rho, Q)$ and $N_{\text{ent}}(\rho, Q)$ are negligible. However, after a time of 1 s, the branches are fully distinguishable and so $N_{\text{ent}}(\rho, Q) \approx N$. We argue that this is the reason why the initial state intuitively (but wrongly) seems to be a macroscopic superposition.

Leggett makes another point which is also relevant here: what if we were to treat individual nucleons as the subsystems? We would find δp to be much larger (as a rough estimate, $\delta x \approx 10^{-15}$ m, so $\delta p \approx \hbar/\delta x \approx 10^{-19}$ kgms $^{-1}$), and N_{ent} would be reduced by another eight orders of magnitude. The lesson here is that the choice of subsystems can have a large impact on N_{ent} .

7.2.2 Molecular diffraction

Now we turn to diffraction experiments of large molecules, using the most recent data involving the largest masses available. Qualitatively, all such experiments operate on the same principle of directing a beam of sufficiently collimated molecules at a grating and observing an interference pattern at some further plane in space. It should be clear that what one demonstrates by observing interference is that the molecule was coherently delocalised over a significant length of the grating covering at least a number of slits. Hence we are interested in the extensive and entangled sizes with respect to Q . Such experiments may operate in the near- or far-field regimes [235]. In the far-field case, we show how a bound on the QFI can be extracted from an interference pattern, while in the more complex near-field case, we provide simple estimates. Regarding the entangled size, we can have in mind the same picture as the above crystal example, and again take the position uncertainty of a single atom in a molecule at rest to be $\delta x \approx 10^{-11}$ m. As long as interference is demonstrated, the coherence length Δx of the molecule is at least the slit separation (typically 10^{-7} m), much larger than δx . So we expect to be in the regime $N_{\text{ent}}(\rho, Q) \approx N$, modulo a reduced visibility factor for which we unfortunately do not have enough information, and will presently ignore.

$|\psi_1\rangle^{\otimes N}$, where $|\langle\psi_0|\psi_1\rangle|^2 = 1 - \epsilon^2$ with small ϵ . It was argued in [234] from two perspectives that the effective size of such a state should be $N\epsilon^2$.

7.2.2.1 Far-field estimates

Suppose the beam of molecules moving along the positive z -direction encounters the grating at $z = 0$, and is detected at $z = L$. Let us write the propagating state of a single molecule as a function of z . As is well-known from classical optics [12], the wavefunction at $z = L$ is related to that at $z = 0$ by a Fourier transform. In one dimension:

$$\psi(x'; L) \propto \int dx \psi(x; 0) e^{-ik_z x x' / L}, \quad (7.22)$$

where k_z is the wavenumber in the z -direction. From this, we see that a phase transformation at the grating $\psi(x; 0) \rightarrow \psi(x; 0) e^{isx/\hbar}$ corresponds to a translation of the interference pattern $\psi(x'; L) \rightarrow \psi(x' - \frac{sL}{\hbar k_z}; L)$. Therefore the QFI with respect to X_{CM} at the grating can be expressed as

$$\mathcal{F}(\rho(0), X_{CM}) = \left(\frac{L}{\hbar k_z} \right)^2 \mathcal{F}(\rho(L), P) \quad (7.23)$$

$$\begin{aligned} \Rightarrow \mathcal{F}(\rho(0), Q) &= \left(\frac{ML}{\hbar k_z} \right)^2 \mathcal{F}(\rho(L), P) \\ &= \left(\frac{L}{v_z} \right)^2 \mathcal{F}(\rho(L), P), \end{aligned} \quad (7.24)$$

where $v_z = \hbar k_z / M$ is the velocity in the z -direction.

Note that $P/\hbar$ is the generator of spatial displacements in the x -direction. So the classical translational Fisher information of the observed probability distribution $p(x)$ at $z = L$ gives a lower bound to the QFI at $z = 0$:

$$\mathcal{F}(\rho(0), Q) \geq \left(\frac{\hbar L}{v_z} \right)^2 \mathcal{F}_c(p(x)). \quad (7.25)$$

This lower bound will be the value that we estimate for the QFI.

A recent experiment in the far-field regime was performed by Juffmann et al. [236] using two types of molecule of mass $514m_u$ and $1298m_u$. The grating has periodicity $d = 100$ nm, the beam has a velocity $v_z \approx 150$ ms⁻¹, and the distance between the grating and detection is $L = 564$ mm. For both molecules, a portion of the interference pattern is given in Figure S2 of [236], along with a (rather accurate) theoretical fit. We have taken points from this curve and smoothly interpolated to obtain a function whose Fisher information can be found numerically. The exact form of the interpolating function will not make a significant difference to the estimates being made here.

For the molecule PcH_2 with $N = 58$, $M = 514m_u$, we find $\mathcal{F}_c(p(x)) \approx 1.3 \times 10^{11} \text{ m}^{-2}$, giving $N_{\text{ext}}(\rho, Q) \approx 6.8 \times 10^{11}$. For $\text{F}_{24}\text{PcH}_2$ with $N = 114$, $M = 1298m_u$, we have $\mathcal{F}_c(p(x)) \approx 7.7 \times 10^9 \text{ m}^{-2}$ and $N_{\text{ext}}(\rho, Q) \approx 3.9 \times 10^{10}$.

It is also interesting to compute an effective “coherence length” at the grating, defined by

$$\chi(\rho) := \frac{\sqrt{\mathcal{F}(\rho, Q)}}{2M}. \quad (7.26)$$

For PcH_2 and $\text{F}_{24}\text{PcH}_2$ respectively, we find $\chi \approx 85 \text{ nm}$ and $\chi \approx 8 \text{ nm}$. These are both smaller than the grating periodicity of 100 nm , expressing the reduced visibility primarily due to the width and finite transverse coherence of the source, as well as a van der Waals interaction between the molecules and the grating [236].

7.2.2.2 Near-field estimates

The most massive object to have been diffracted so far is the “L12” molecule (chemical formula $\text{C}_{284}\text{H}_{190}\text{F}_{320}\text{N}_4\text{S}_{12}$) of [237], with $N = 810$ and $M = 10123m_u$. This experiment operates in the near-field regime, so there is no simple connection between $\mathcal{F}(\rho, Q)$ at the grating and $\mathcal{F}_c(p(x))$. Instead, we make the crude estimate that the molecule has a transverse coherence length of the same approximate size as the grating periodicity, which here is $d \approx 266 \text{ nm}$. Then $\mathcal{F}(\rho, Q)/4 \approx (Md)^2$ and $N_{\text{ext}}(\rho, Q) \approx 2.6 \times 10^{15}$.

If χ is, say, an of magnitude smaller than d , then we would have $N_{\text{ext}}(\rho, Q) \sim 10^{13}$. A more precise estimate is not currently possible with the available data and techniques.

Along the same lines, we also include an estimate from a typical neutron diffraction experiment [238] for comparison. Here, the slit separation is $d = 107 \mu\text{m}$. Again assuming unit visibility, we have $N_{\text{ext}}(\rho, Q) \approx (Md/2Q_0)^2 \approx 1.0 \times 10^{12}$. Clearly, $N_{\text{ent}}(\rho, Q) \approx 1$ with a single subsystem.

7.2.3 Mechanical oscillators

The centre-of-mass motion of a mechanical oscillator is well described as a quantum harmonic oscillator of some effective mass M and resonance frequency ω . The zero-point motion of the ground state entails a position uncertainty of $\Delta X = \sqrt{\frac{\hbar}{2M\omega}}$ and a total momentum uncertainty of $\Delta P = \sqrt{\frac{\hbar M\omega}{2}}$ [64]. For a thermal state, the QFI is

reduced by a factor depending on the temperature (see Appendix E for the proof):

$$\frac{1}{4}\mathcal{F}(\rho, Q) = \frac{M^2\Delta X^2}{1+2\bar{n}}, \quad \frac{1}{4}\mathcal{F}(\rho, P) = \frac{\Delta P^2}{1+2\bar{n}}, \quad (7.27)$$

where $\bar{n} = \text{Tr}(\rho\hat{n}) = (e^{\beta\hbar\omega} - 1)^{-1}$ is the expected number of vibrational quanta. The extensive sizes in Q and P follow by normalising by Q_0^2 and P_0^2 respectively.

For the entangled size, we need the single-particle variances. As for the crystal example, we estimate the position uncertainty of each atom to be $\delta x \approx 10^{-11}$ m, whenever the whole object is at rest. The variance from the centre-of-mass motion is $(\Delta X)^2(1+2\bar{n})$ (see Appendix E). Adding these independent contributions gives $V(\rho_1, Q) \approx m^2[(\delta x)^2 + (\Delta X)^2(1+2\bar{n})]$. The same logic applies to P , using $\delta p \approx \frac{\hbar}{2\delta x}$ and $\Delta p = \Delta P/N$. Thus the estimated entangled sizes are

$$N_{\text{ent}}(\rho, Q) = \frac{Nr_q^2}{f(1+fr_q^2)}, \quad N_{\text{ent}}(\rho, P) = \frac{Nr_p^2}{f(1+fr_p^2)}, \quad (7.28)$$

where $r_q = \Delta X/\delta x$, $r_p = \Delta p/\delta p$ and $f := 1+2\bar{n}$.

In Table 7.1 we present a survey of these numbers for some recent mechanical oscillator experiments. In each case, we take m to be an average of all of the atomic masses in the oscillator. Note that we have not calculated values for $N_{\text{ent}}(\rho, P)$. This is because the ratio r_p , which can be expressed as $\frac{\delta x}{N\Delta X}$, is at most of order 10^{-7} in each experiment, making the entangled size in P negligible. In the most recent experiment presented here [239], which is also the most massive, no attempt at cooling was made and so the large $\bar{n}$ is due to room temperature. We also include two proposed experiments.

In the first [240], although the mass is not as great as in previous experiments, the smaller value of ω would result in larger extensive and entangled sizes with respect to Q . We present ideal numbers for zero temperature – it is argued in [240] that cooling close to the ground state should be feasible. The second proposal [241] is for a scheme to test quantum gravity models predicting alterations to the canonical commutator. We take the least demanding parameters required in that work, namely the smallest mass and highest tolerable temperature.

7.2.4 Proposed quantum gravity test

We now include an estimate for a recent proposal to test the non-classicality of gravity by witnessing its ability to entangle two test masses [245]. This requires each test

Ref.	m (m_u)	$\omega/2\pi$ (Hz)	$\bar{n}$	N	$N_{\text{ext}}(\rho, Q)$	$N_{\text{ext}}(\rho, P)$	$N_{\text{ent}}(\rho, Q)$
[242]	27	1.06×10^7	0.34	1.1×10^{12}	3×10^{18}	5×10^7	1×10^5
[243]	28	3.90×10^9	0.84	6.7×10^9	3×10^{13}	8×10^7	2×10^2
[244]	20	7.80×10^7	1.7	6.0×10^{13}	6×10^{18}	6×10^9	7×10^3
[239]	20	1.06×10^5	6×10^7	3.0×10^{15}	9×10^{15}	1×10^1	1×10^{-1}
[240]	20	5.0×10^2	≈ 0	1.5×10^{14}	1×10^{25}	4×10^5	5×10^9
[241]	20	1.0×10^5	30	3.0×10^{14}	2×10^{21}	3×10^6	4×10^5

Table 7.1: Experimental parameters and estimated extensive and entangled sizes for six different mechanical oscillators, the last two being proposals.

mass of $M \approx 10^{-14}$ kg to be initially put into a spatial superposition state where the branches are separated by $\Delta X \approx 250 \mu\text{m}$. As for the above cases, we assume that each atom has a position uncertainty of $\delta x \approx 10^{-11}$ m when the object is at rest. Then for one test mass we have $N_{\text{ext}}(\rho, Q) \approx 8 \times 10^{38}$, $N_{\text{ent}}(\rho, Q) \approx N \approx 10^{15}$, without decoherence.

7.2.5 Hypothetical Schrödinger’s cat

For comparison with a truly macroscopic system, we estimate the extensive and entangled sizes of a hypothetical cat in a superposition state as envisaged by Schrödinger. The “dead” and “alive” branches contain an object of mass of order $M \approx 1$ kg, corresponding to $N \approx 10^{26}$ atoms, with a difference in position of around $\Delta X \approx 0.1$ m. Then $N_{\text{ext}}(\rho, Q) \approx 10^{72}$ and $N_{\text{ent}}(\rho, Q) \approx N \approx 10^{26}$.

7.3 Conclusions

In Fig. 7.1 we have plotted extensive and entangled sizes for the systems studied above. It is clear that the recent experiments with mechanical oscillators are capable of leading in both measures over diffracting molecules. As noted above, for mechanical oscillators, both measures are small with respect to P . Instead for Q , we typically find an intermediate regime where $1 \ll N_{\text{ent}}(\rho, Q) \ll N$, indicating a degree of similarity to a cat state in position space. In this regime, it is easily shown that $N_{\text{ent}}(\rho, Q) \approx \hbar/(2m\omega(\delta x)^2)$, demonstrating that smaller frequencies ω provide the path to larger entangled sizes (since m and δx do not vary greatly). Physically, this corresponds to weaker trapping in position, so the ground state is delocalised over a greater distance. The extensive size is similarly $N_{\text{ext}}(\rho, Q) \approx (M\hbar)/(2\omega)$ – also increasing with the

total mass. More precise estimates for future experiments could be obtained from tomography via the Wigner function (a technique for such a measurement was shown in [239]).

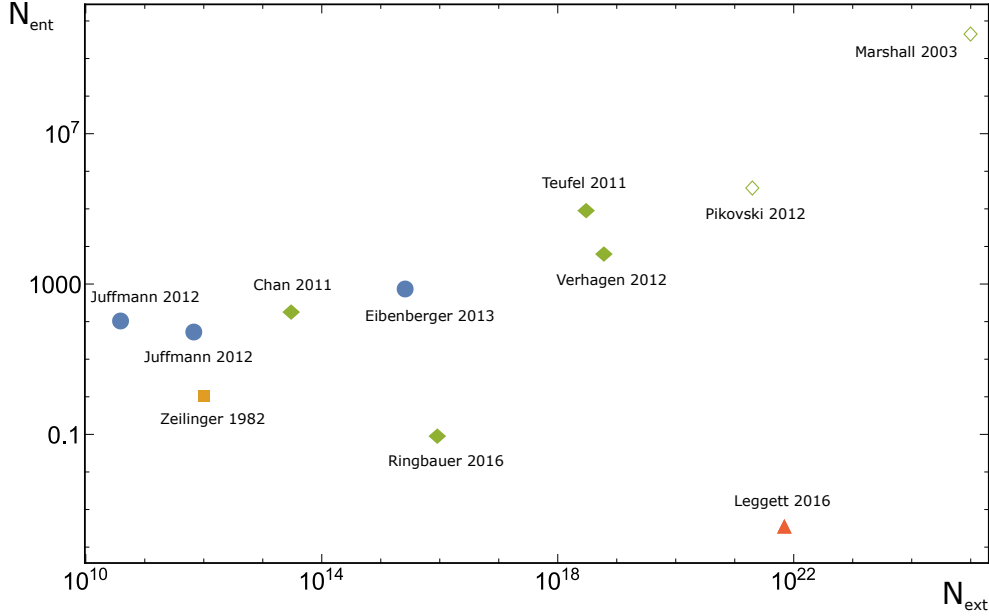


Figure 7.1: Extensive and entangled sizes N_{ext} , N_{ent} (showing the maximum over Q and P) for the following systems: molecular diffraction [236, 237] (blue circles); neutron diffraction [238] (orange square); mechanical oscillators [239, 242–244] (green diamonds) including proposals [240, 241] (green unfilled diamonds); Leggett’s crystal thought experiment [231] (red triangle).

For diffracting molecules, a more precise estimation method with fewer assumptions is needed in the near-field case. Perhaps this could be attempted by a detailed theoretical model of the system, but it is not presently clear how a more direct calculation from experimental data could be done.

There are other systems often described as macroscopic quantum phenomena which remain to be analysed. Of particular note are superconducting qubits [59], which have been the subject of much debate in the macroscopicity literature [65, 66, 231, 232]. These come in a number of varieties [59]; in the earliest works showing strong experimental evidence for superpositions of mesoscopically distinct states in such systems, the two branches contain different currents [60, 61]. On one hand, it is simple to estimate an extensive size for these systems: the total momentum (integrated along the circuit) in a definite-current branch is $P = m_e l I / e$, where m_e and e are the mass and charge of the electron, l is the length of the circuit, and I is the

current. Assuming a pure state, the momentum-space extensive size is

$$N_{\text{ext}}(\rho, P) = \left(\frac{m_e l \Delta I}{2eP_0} \right)^2, \quad (7.29)$$

where ΔI is the difference in current between the branches. From [61], we have $\Delta I \approx 2 \mu\text{A}$, $l = 560 \mu\text{m}$, giving $N_{\text{ext}}(\rho, P) \approx 2.5 \times 10^6$. This is comparable to the momentum-space extensive sizes of some of the above systems, but much less than their position-space sizes. However, it is uncertain how to calculate an appropriate entangled size, as the choice of subsystems is not obvious. This is a point of contention: whereas Korsbakken et al. [232, 246] talk about single-electron modes, Leggett [231] argues that Cooper pairs of electrons are the relevant subsystems. We do not attempt to settle this issue here, but leave it for future work.

Furthermore, Bose-Einstein condensates are often described as macroscopic quantum states [247]. While a detailed analysis remains to be done, we fully expect that the entangled size of such states is small. As others have observed [65, 248], the properties of such condensates are well explained by only two-body correlations. It will also be interesting to look at condensates which are split between two locations while maintaining phase coherence [249]. Nevertheless, some fundamental questions have to be answered, such as whether it makes sense to use atoms as subsystems, as here they are indistinguishable particles.

Chapter 8

Conclusions and outlook

In this thesis, we have explored a variety of ways of formulating a resource theory of quantum coherence. In Chapter 3, we examined the foundations of the resource theory of speakable coherence, and found that restricting the almost canonical BCP set of incoherent operations to SI operations enables a physical picture to be given by a dilation involving an ancilla. We argue that this makes the SI set the best physically motivated set for speakable coherence, since it takes into account the coherent resources needed to perform a general non-unitary incoherent operation. Most results in the literature about state transformations use the BCP set, and so there are now open questions to be resolved about the operational differences of the SI set, as outlined at the end of Chapter 3. One feature of the BCP set, that is lost upon restricting to the SI set, is the ability to use a maximally coherent state to generate an arbitrary unitary operation on another state. The Kraus operators shown to accomplish this in [40] are explicitly non-SI. The theory of asymmetry is similar to the BCP theory in this way – states acting as reference frames can be used to perform asymmetric operations [141]. In the SI theory, measurements in a coherent basis are not treated as free, and such operations are not facilitated by the addition of coherent ancilla states. This suggests that, in the resource theory of speakable coherence, coherent-basis measurements need to be treated as resources which are independent of coherent state resources (i.e., the two are not mutually convertible). A full exploration of these ideas remains to be done.

Many questions remain about the full scope of speakable coherence as a useful resource in physical or computational settings. Coherence is obviously necessary for a quantum advantage, otherwise the dynamics are describable by a classical probabilis-

tic model. However, it is less trivial to determine when coherence is also sufficient. An example of this subtlety is described in [41]: the state injection model of quantum computation [250]. Here, the only permitted gates are the Hadamard, C-NOT and a particular phase gate. The Gottesmann-Knill theorem [72] says that the evolution of certain states (known as stabiliser states) under such gates can be efficiently simulated on a classical computer – even though the coherence may be arbitrarily large. This shows that coherence is not sufficient for a speed-up in this particular model. Outside of computation, the role of coherence in thermodynamics and transport phenomena remain as open projects in their own right; so far, no concrete link has been given between these areas and the free operations of speakable coherence. It seems that such a link would require a physical justification of the free unitaries (in particular, arbitrary permutations of the incoherent basis) as being “easy” operations.

Next, in Chapter 4, we examined the connections between coherence and quantum correlations, specifically of the discord type. The main initial result was that the amount of discord created by incoherent operations is upper bounded by the coherence initially present. Moreover, the consumption of coherence provides a bound in the case of discord defined in the manner of Ollivier and Zurek [42]. This is to some extent a generalisation of previous results linking coherence and entanglement [138]. We gave evidence to suggest that the consumption of coherence was a better indication than creation of discord for the speed-up of the DQC1 algorithm. This is still an unsettled claim, mainly because it is unclear how to fairly compare the quantum algorithm with a classical version.

The second part of the chapter focused on basis-dependent discord. This is something that has been avoided in past work on discord; with light shed on coherence as a resource, we can now appreciate the relevance of basis-dependent quantities. We have characterised BDD to show that it can indeed be interpreted as the quantumness of correlations with respect to a fixed basis. Unlike the basis-independent case, it was also possible to say something about the behaviour of BDD under local incoherent operations. Failure of monotonicity can be associated with the loss of a classical memory. This might give an indication of how to formulate a resource theory of BDD: tracing out a classical memory must *not* be treated as a free operation. Perhaps this can be motivated in a quantum cryptography situation where a party is unable to guarantee that their classical memory is kept hidden from an eavesdropper.

If a sensible resource theory of BDD can be found, this may pave the way for a full resource theory of discord.

In the remaining sections of the thesis, we employed a theory of unspeakable coherence (a type of asymmetry) as a framework for characterising macroscopic superposition states. We argued why this is an appropriate formalism, and showed that some previously proposed measures of macroscopicity can be understood as measuring coherence at their core. The monotonicity criterion also enabled us to point out a flaw in one of these measures. We showed that the quantum Fisher information is a good candidate for a macroscopicity measure, as it is the asymptotic formation rate of this resource theory. A natural question is whether the asymptotic cost can be obtained by regularisation, as works for entanglement [251] – this would again give the QFI, by additivity. As we have discussed, distillation may be a more complex issue.

In Chapter 6, we demonstrated how the concept of macroscopic distinguishability can be formalised into measures which obey the resource theory criteria. These are based on the branch-distinguishing power of a measurement device with finite resolution. By thinking of the environment in a dephasing process as such a measurement device, we showed a direct connection between the coherence measures and the fragility of micro-macro entangled states. The greater the coherence, the more fragile the entanglement – this shows one way in which decoherence leads to an effective quantum-to-classical transition. Furthermore, the intrinsic uncertainty of the measurement device was found to act as a parameter which interpolates between a measure of speakable coherence (in the high precision limit) and a measure suitable for macroscopicity (in the low precision limit). We also found that the QFI is recovered in the latter limit. The most interesting prospect for future work here is to generalise the measurement model and thereby derive results about fragility under a larger class of decoherence models.

Finally, in Chapter 7, we applied the framework to evaluate the macroscopicity of past, proposed and hypothetical experiments with mechanical systems. Similarly to Leggett’s ideas, we argued that macroscopicity may not be given by a single number. Instead, we proposed two measures based on the QFI. The first, the extensive size, calculates the total amount of coherence in a system, and manages to capture the “spread” of coherence with respect to position or momentum. However, it is additive

with respect to subsystems, so on its own could be misleading for comparing systems with different numbers of particles. Therefore we also introduced the entangled size, which describes how coherence is shared between subsystems and acts as a multipartite entanglement witness. From experimental estimates, we found that recent micro-mechanical systems outperformed diffracting molecules in both measures. We plan to extend this analysis to more systems, including spins and photons, for which similar initial work has been done recently [223], as well as quantum condensates. Some thought needs to be put into whether such qualitatively different systems can really be compared, and in addition how to sensibly define subsystems in cases of indistinguishable particles.

It is also worth exploring different ways in which the QFI can be measured more directly in experiments. Recent developments include lower bounds where the number of required measurements scales exponentially better than tomography [5, 7, 201], a connection to dynamic susceptibilities (which are routinely measured in condensed matter physics by neutron scattering) [252], and a technique for overcoming finite measurement precision [253].

Let us end with a comment about coherence in CV phase space. The approach we have taken here is to quantify macroscopicity in phase space by maximising the coherence (QFI) over quadratures. We saw that this evaluated optical coherent states as the most classical pure states, fitting with expectation. However, one might ask for a more direct approach with a resource theory where optical coherent states and their classical mixtures comprise the free states. Some ideas along these lines have been presented in, for example [254]. The relevant free operations need to be determined; we can expect this set to at least include displacements and rotations of phase space. This type of project could lead to a full resource theory of non-classicality in optics. It may also be possible to generalise this to other systems in which there is a natural set of classical states but no single preferred basis.

Appendix A

Generalised QFI

A.1 Background

See [202, 219] for a detailed introduction; we follow the notation of [202].

The set of symmetric, normalised **operator monotone functions** $\mathcal{F}_{\text{op}}$ consists of $f : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ such that

- i) For any hermitian operators A, B such that $0 \leq A \leq B$, we have $0 \leq f(A) \leq f(B)$;
- ii) $f(x) = xf(x^{-1})$;
- iii) $f(1) = 1$.

These functions can be put into one-to-one correspondence with the set $\mathcal{M}_{\text{op}}$ of **matrix means** $m(A, B)$; see [202] for a list of defining properties. The link between the two sets is

$$m_f(A, B) := A^{\frac{1}{2}} f(A^{-\frac{1}{2}} B A^{-\frac{1}{2}}) A^{\frac{1}{2}}, \quad (\text{A.1})$$

which reduces to $m_f(A, B) = Af(BA^{-1})$ for commuting A, B .

Operator monotone functions and matrix means also have a bijection with the set of **monotone metrics** which give rise to norms $\|X\|_{\rho, f}$ defined by

$$\|X\|_{\rho, f}^2 := \text{Tr} \left(X m_f(L_\rho, R_\rho)^{-1}(X) \right), \quad (\text{A.2})$$

where R_ρ and L_ρ are the right- and left-multiplication superoperators: $R_\rho(X) = X\rho$, $L_\rho(X) = \rho X$. The monotonicity property of these metrics is that, for any channel $\mathcal{E}$, one has

$$\|\mathcal{E}(X)\|_{\mathcal{E}(\rho), f} \leq \|X\|_{\rho, f}. \quad (\text{A.3})$$

Now the QFI associated with $f \in \mathcal{F}_{\text{op}}$ is

$$I_f(\rho, A) := \frac{f(0)}{2} \| -i[A, \rho] \|_{\rho, f}^2, \quad (\text{A.4})$$

which is normalised such that for pure states, $I_f(|\psi\rangle\langle\psi|, A) = V(|\psi\rangle\langle\psi|, A) = \langle\psi|A^2|\psi\rangle - \langle\psi|A|\psi\rangle^2$. For an arbitrary state with a spectral decomposition $\rho = \sum_i \lambda_i |\psi_i\rangle\langle\psi_i|$, it can be shown [198] that

$$I_f(\rho, A) = \frac{f(0)}{2} \sum_{i,j} \frac{(\lambda_i - \lambda_j)^2}{\lambda_j f(\lambda_i/\lambda_j)} |\langle\psi_i|A|\psi_j\rangle|^2, \quad (\text{A.5})$$

where each term in the sum is taken to be zero whenever $\lambda_i = \lambda_j$.

We now give a fact which will be useful later.

Lemma 11. *For any set of states ρ_μ and normalised probabilities p_μ , and an orthonormal set $\{|\mu\rangle\}$,*

$$I_f(\sum_\mu p_\mu \rho_\mu \otimes |\mu\rangle\langle\mu|, A \otimes I) = \sum_\mu p_\mu I_f(\rho_\mu, A). \quad (\text{A.6})$$

Proof. Let each ρ_μ have a spectral decomposition $\rho_\mu = \sum_i \lambda_{i|\mu} |\psi_{\mu,i}\rangle\langle\psi_{\mu,i}|$. Using (A.5) and defining $\lambda_{\mu,i} := p_\mu \lambda_{i|\mu}$, we then have

$$\begin{aligned} I_f(\sum_\mu p_\mu \rho_\mu \otimes |\mu\rangle\langle\mu|, A \otimes I) &= \frac{f(0)}{2} \sum_{\mu, \nu, i, j} \frac{(\lambda_{\mu,i} - \lambda_{\nu,j})^2}{\lambda_{\nu,j} f(\lambda_{\mu,i}/\lambda_{\nu,j})} |\langle\psi_{\mu,i}|\langle\mu|(A \otimes I)|\psi_{\nu,j}\rangle|\nu\rangle|^2 \\ &= \frac{f(0)}{2} \sum_{\mu, i, j} \frac{(\lambda_{\mu,i} - \lambda_{\mu,j})^2}{\lambda_{\mu,j} f(\lambda_{\mu,i}/\lambda_{\mu,j})} |\langle\psi_{\mu,i}|A|\psi_{\mu,j}\rangle|^2 \\ &= \frac{f(0)}{2} \sum_{\mu, i, j} \frac{p_\mu^2 (\lambda_{i|\mu} - \lambda_{j|\mu})^2}{p_\mu \lambda_{j|\mu} f(\lambda_{i|\mu}/\lambda_{j|\mu})} |\langle\psi_{\mu,i}|A|\psi_{\mu,j}\rangle|^2 \\ &= \sum_\mu p_\mu I_f(\rho_\mu, A). \end{aligned} \quad (\text{A.7})$$

□

Corollary 1. *$I_f(\rho, A)$ is convex in ρ .*

Proof. This follows from Lemma 11 by tracing out the ancillary system, as I_f must be monotonically decreasing under partial trace:

$$\sum_\mu p_\mu I_f(\rho_\mu, A) = I_f(\sum_\mu p_\mu \rho_\mu \otimes |\mu\rangle\langle\mu|) \geq I_f(\sum_\mu p_\mu \rho_\mu, A). \quad (\text{A.8})$$

□

A.2 Deterministic monotone

The unitary case was shown already in [202]. Noting that $-i[A, \rho] = \partial_t \mathcal{U}_t(\rho)|_{t=0}$, we have from (A.4) that

$$I_f(\rho, A) = \frac{f(0)}{2} \|\partial_t \mathcal{U}_t(\rho)|_0\|_{\rho, f}^2. \quad (\text{A.9})$$

The linearity of $\mathcal{E}$ and the monotonicity property (A.3) then give

$$\begin{aligned} \|\partial_t \mathcal{U}_t(\mathcal{E}(\rho))\|_{\mathcal{E}(\rho), f} &= \|\partial_t \mathcal{E}(\mathcal{U}_t(\rho))\|_{\mathcal{E}(\rho), f} \\ &= \|\mathcal{E}(\partial_t \mathcal{U}_t(\rho))\|_{\mathcal{E}(\rho), f} \\ &\leq \|\partial_t \mathcal{U}_t(\rho)\|_{\rho, f}, \end{aligned} \quad (\text{A.10})$$

so that $I_f(\mathcal{E}(\rho), A) \leq I_f(\rho, A)$.

A.3 Ensemble monotone

In this case, we now have a quantum instrument defined by a set of TI maps $\{\mathcal{E}_\mu\}$ which are not necessarily trace preserving, while the sum $\sum_\mu \mathcal{E}_\mu$ is. For every quantum instrument, one can construct a channel (i.e., so that trace is preserved) by including in the output an ancilla that records which outcome was obtained via a set of orthonormal states $\{|\mu\rangle\}$:

$$\mathcal{E}'(\rho) := \sum_\mu \mathcal{E}_\mu(\rho) \otimes |\mu\rangle\langle\mu|. \quad (\text{A.11})$$

Tracing out the ancilla results in the channel $\sum_\mu \mathcal{E}_\mu$. It is clear that $\mathcal{E}'$ is TI whenever each of the $\mathcal{E}_\mu$ is. Writing $\mathcal{E}'(\rho) = \sum_\mu p_\mu \rho_\mu \otimes |\mu\rangle\langle\mu|$, we can apply Lemma 11 followed by deterministic monotonicity to find

$$\begin{aligned} \sum_\mu p_\mu I_f(\rho_\mu, A) &= I_f\left(\sum_\mu p_\mu \rho_\mu \otimes |\mu\rangle\langle\mu|, A \otimes I\right) \\ &= I_f(\mathcal{E}'(\rho), A \otimes I) \\ &\leq I_f(\rho, A). \end{aligned} \quad (\text{A.12})$$

Appendix B

N_{LJ} decoherence model

We first assemble some useful facts for a single mode [76]:

$$\langle \beta | \alpha \rangle = e^{-\frac{1}{2}|\beta-\alpha|^2 + \frac{1}{2}(\alpha\beta^* - \alpha^*\beta)}, \quad (\text{B.1})$$

$$D(\alpha)D(\beta) = e^{\frac{1}{2}(\alpha\beta^* - \alpha^*\beta)} D(\alpha + \beta), \quad (\text{B.2})$$

$$\frac{1}{\pi} \int d^2\alpha |\alpha\rangle\langle\alpha| = I. \quad (\text{B.3})$$

Using these, we have

$$\begin{aligned} \text{Tr } D(\alpha) &= \frac{1}{\pi} \int d^2\beta \langle \beta | D(\alpha) | \beta \rangle \\ &= \frac{1}{\pi} e^{-|\alpha|^2} \int d^2\beta e^{\alpha\beta^* - \alpha^*\beta} \\ &= \pi \delta^2(\alpha), \end{aligned} \quad (\text{B.4})$$

from which it follows that we can write (in the N -mode case)

$$\rho = \frac{1}{\pi^N} \int d^{2N}\boldsymbol{\alpha} \chi_\rho(\boldsymbol{\alpha}) D(-\boldsymbol{\alpha}). \quad (\text{B.5})$$

Under the model in Result 16(a), using $[x_i, D(\boldsymbol{\alpha})] = \sqrt{2}\text{Re}(\alpha_i)D(\boldsymbol{\alpha})$ and $[p_i, D(\boldsymbol{\alpha})] = \sqrt{2}\text{Im}(\alpha_i)D(\boldsymbol{\alpha})$, we have

$$\begin{aligned} \partial_t \rho &= -\frac{1}{4} \sum_{i=1}^N \frac{1}{\pi^N} \int d^{2N}\boldsymbol{\alpha} (2\text{Re}(\alpha_i)^2 + 2\text{Im}(\alpha_i)^2) \chi_\rho(\boldsymbol{\alpha}) D(-\boldsymbol{\alpha}) \\ &= -\frac{1}{2\pi^N} \int d^{2N}\boldsymbol{\alpha} |\boldsymbol{\alpha}|^2 \chi_\rho(\boldsymbol{\alpha}) D(-\boldsymbol{\alpha}), \end{aligned} \quad (\text{B.6})$$

and so $\partial_t \chi_\rho(\boldsymbol{\alpha}) = -\frac{1}{2}|\boldsymbol{\alpha}|^2 \chi_\rho(\boldsymbol{\alpha})$.

Finally, we calculate the rate of loss of purity:

$$\begin{aligned}
-\frac{1}{2}\partial_t \text{Tr}(\rho^2) &= -\text{Tr}(\rho \partial_t \rho) \\
&= -\frac{1}{\pi^{2N}} \int d^{2N}\boldsymbol{\alpha} d^{2N}\boldsymbol{\beta} \chi_\rho(\boldsymbol{\alpha}) \partial_t \chi_\rho(\boldsymbol{\beta}) \text{Tr}[D(-\boldsymbol{\alpha})D(-\boldsymbol{\beta})] \\
&= -\frac{1}{\pi^{2N}} \int d^{2N}\boldsymbol{\alpha} d^{2N}\boldsymbol{\beta} \chi_\rho(\boldsymbol{\alpha})^* \partial_t \chi_\rho(\boldsymbol{\beta}) \text{Tr}[D(\boldsymbol{\alpha})D(-\boldsymbol{\beta})] \\
&= -\frac{1}{\pi^{2N}} \int d^{2N}\boldsymbol{\alpha} d^{2N}\boldsymbol{\beta} \chi_\rho(\boldsymbol{\alpha})^* \partial_t \chi_\rho(\boldsymbol{\beta}) e^{\frac{1}{2}(-\boldsymbol{\alpha} \cdot \boldsymbol{\beta}^* + \boldsymbol{\alpha}^* \cdot \boldsymbol{\beta})} \text{Tr}[D(\boldsymbol{\alpha} - \boldsymbol{\beta})] \\
&= -\frac{1}{\pi^{2N}} \int d^{2N}\boldsymbol{\alpha} d^{2N}\boldsymbol{\beta} \chi_\rho(\boldsymbol{\alpha})^* \partial_t \chi_\rho(\boldsymbol{\beta}) e^{\frac{1}{2}(-\boldsymbol{\alpha} \cdot \boldsymbol{\beta}^* + \boldsymbol{\alpha}^* \cdot \boldsymbol{\beta})} \pi^N \delta^{2N}(\boldsymbol{\alpha} - \boldsymbol{\beta}) \\
&= -\frac{1}{\pi^N} \int d^{2N}\boldsymbol{\alpha} \chi_\rho(\boldsymbol{\alpha})^* \partial_t \chi_\rho(\boldsymbol{\alpha}) \\
&= \frac{1}{2\pi^N} \int d^{2N}\boldsymbol{\alpha} |\boldsymbol{\alpha}|^2 |\chi_\rho(\boldsymbol{\alpha})|^2 \\
&= N_{LJ}(\rho)
\end{aligned} \tag{B.7}$$

as required.

Appendix C

Details of monotonicity of dist_Δ

We write $|\psi\rangle = \sum_i \psi_i |i\rangle$, $|\phi_\mu\rangle = \sum_i \phi_i^\mu |i\rangle$. Then

$$\begin{aligned} \sum_\mu w_\mu p_a^\mu &= \sum_\mu \langle \psi | K_\mu^\dagger \Pi_{a+\delta_\mu} K_\mu | \psi \rangle \\ &= \sum_\mu \sum_{i: a_i=a+\delta_\mu} \sum_{j,k} \langle \psi | K_\mu^\dagger | i \rangle \langle i | K_\mu | \psi \rangle \\ &= \sum_\mu \sum_{i: a_i=a+\delta_\mu} \sum_{j,k} \psi_j^* \psi_k c_{ij}^{\mu *} c_{ik}^\mu. \end{aligned} \tag{C.1}$$

The condition $I = \sum_\mu K_\mu^\dagger K_\mu$ implies (taking the $\langle j | | k \rangle$ matrix element) that

$$\begin{aligned} \delta_{j,k} &= \sum_\mu \sum_i \langle j | K_\mu^\dagger | i \rangle \langle i | K_\mu | k \rangle \\ &= \sum_\mu \sum_i c_{ij}^{\mu *} c_{ik}^\mu \\ &= \sum_\mu \sum_{i: a_i-a_j=\delta_\mu} c_{ij}^{\mu *} c_{ik}^\mu, \end{aligned} \tag{C.2}$$

where the last line follows from the form of the Kraus operators (3.34). Putting this into the previous equation gives

$$\begin{aligned} \sum_\mu w_\mu p_a^\mu &= \sum_{j,k: a_j=a} \delta_{j,k} \psi_j^* \psi_k \\ &= \sum_{j: a_j=a} |\psi_j|^2 \\ &= p_a. \end{aligned} \tag{C.3}$$

Appendix D

Degradation of entanglement of formation

Let us consider interacting S with the measurement device (or environment) M , and then projecting M onto an eigenstate of the pointer observable x_M . This occurs with probability $p_M(x)$ and results in

$$p_M(x)\rho_{mS|x} := \langle x |_M e^{-iA_S p_M} (\rho_{mS} \otimes |\chi_\Delta\rangle\langle\chi_\Delta|_M) e^{iA_S p_M} |x\rangle_M. \quad (\text{D.1})$$

Now, since ρ_{mS} is pure, so is each $\rho_{mS|x}$. Hence a possible pure state ensemble for the final state is $\rho'_{mS} = \int dx p_M(x)\rho_{mS|x}$, and therefore

$$E_f(\rho'_{mS}) \leq \int dx p_M(x) E_E(\rho_{mS|x}) = \int dx p_M(x) S(\rho_{S|x}), \quad (\text{D.2})$$

where $\rho_{S|x} = \text{Tr}_m \rho_{mS|x}$. Observe that the right hand side of the above equation is exactly the conditional entropy $H(B|A)$ between Alice and Bob in the set-up defining dist_Δ (Definition 19). This can be expressed as $H(B) - J(A : B)$ and $H(B) = H(\{p_i\})$. Therefore

$$E_f(\rho'_{mS}) \leq H(\{p_i\}) - \text{dist}_\Delta(\rho_{mS}, A_S), \quad (\text{D.3})$$

and furthermore $H(\{p_i\}) = S(\rho_S) = E_f(\rho_{mS})$.

Appendix E

Quadrature QFI and variance in a thermal state

A thermal state of a single mode can be written as $\rho = \sum_{n=0}^{\infty} (1-p)p^n |n\rangle\langle n|$, where $p = e^{-\beta\hbar\omega}$. The expression (5.7) for the QFI then gives

$$\mathcal{F}(\rho, x) = 4(1-p) \sum_{n=0}^{\infty} \sum_{m=0}^{n-1} \frac{(p^n - p^m)^2}{p^n + p^m} |\langle m|x|n\rangle|^2. \quad (\text{E.1})$$

From the matrix element $\langle m|X|n\rangle = \Delta X \sqrt{n} \delta_{n,m+1}$ for $n > m$, we obtain

$$\mathcal{F}(\rho, x) = \frac{4\Delta X^2(1-p)^3}{1+p} \sum_{n=0}^{\infty} (n+1)p^n \quad (\text{E.2})$$

$$= 4\Delta X^2 \left(\frac{1-p}{1+p} \right) = \frac{4\Delta X^2}{1+2\bar{n}}. \quad (\text{E.3})$$

Similarly, the variance can be calculated using $\langle n|x|n\rangle = 0$, $\langle n|x^2|n\rangle = \Delta X^2(1+2n)$, so

$$V(\rho, x) = \text{Tr}(\rho x^2) = \Delta X^2(1+2\bar{n}). \quad (\text{E.4})$$

Bibliography

- [1] Yadin, B. & Vedral, V. Quantum macroscopicity versus distillation of macroscopic superpositions. *Phys. Rev. A* **92**, 022356 (2015).
- [2] Yadin, B. & Vedral, V. General framework for quantum macroscopicity in terms of coherence. *Phys. Rev. A* **93**, 022122 (2016).
- [3] Yadin, B., Ma, J., Girolami, D., Gu, M. & Vedral, V. Quantum Processes Which Do Not Use Coherence. *Phys. Rev. X* **6**, 041028 (2016).
- [4] Ma, J., Yadin, B., Girolami, D., Vedral, V. & Gu, M. Converting Coherence to Quantum Correlations. *Phys. Rev. Lett.* **116**, 160407 (2016).
- [5] Girolami, D. & Yadin, B. Witnessing Multipartite Entanglement by Detecting Asymmetry. *Entropy* **19**, 124 (2017).
- [6] Sekatski, P. *et al.* General measure for macroscopic quantum states beyond “dead and alive”. *arXiv:1704.02270* (2017).
- [7] Zhang, C. *et al.* Determining the speed of multipartite quantum systems by few local measurements. *arXiv:1611.02004* (2016).
- [8] Einstein, A. Über einen die Erzeugung und Verwandlung des Lichtes betreffenden heuristischen Gesichtspunkt. *Ann. Phys.* **322**, 132–148 (1905).
- [9] Bohr, N. I. On the constitution of atoms and molecules. *Philos. Mag. Ser. 6* **26**, 1–25 (1913).
- [10] Sommerfeld, A. Zur Quantentheorie der Spektrallinien. *Ann. Phys.* **356**, 1–94 (1916).

- [11] Davisson, C. J. & Germer, L. H. Reflection of Electrons by a Crystal of Nickel. *Proc. Natl. Acad. Sci.* **14**, 317–322 (1928).
- [12] Brooker, G. *Modern Classical Optics* (Oxford University Press, 2003).
- [13] Glauber, R. J. The Quantum Theory of Optical Coherence. *Phys. Rev.* **130**, 2529–2539 (1963).
- [14] Sudarshan, E. C. G. Equivalence of Semiclassical and Quantum Mechanical Descriptions of Statistical Light Beams. *Phys. Rev. Lett.* **10**, 277–279 (1963).
- [15] Streltsov, A., Adesso, G. & Plenio, M. B. Quantum Coherence as a Resource. *arXiv:1609.02439* (2016).
- [16] Feynman, R. P. Simulating physics with computers. *Int. J. Theor. Phys.* **21**, 467–488 (1982).
- [17] Georgescu, I. M., Ashhab, S. & Nori, F. Quantum simulation. *Rev. Mod. Phys.* **86**, 153–185 (2014).
- [18] Deutsch, D. Quantum Theory, the Church-Turing Principle and the Universal Quantum Computer. *Proc. R. Soc. A Math. Phys. Eng. Sci.* **400**, 97–117 (1985).
- [19] Montanaro, A. Quantum algorithms: an overview. *npj Quantum Inf.* **2**, 15023 (2016).
- [20] Ladd, T. D. *et al.* Quantum computers. *Nature* **464**, 45–53 (2010).
- [21] Bennett, C. H. & Brassard, G. Quantum cryptography: Public key distribution and coin tossing. In *Proc. IEEE Int. Conf. Comput. Syst. Signal Process.*, 175–179 (1984).
- [22] Ekert, A. K. Quantum cryptography based on Bell’s theorem. *Phys. Rev. Lett.* **67**, 661–663 (1991).
- [23] Scarani, V. *et al.* The security of practical quantum key distribution. *Rev. Mod. Phys.* **81**, 1301–1350 (2009).
- [24] Gisin, N. & Thew, R. Quantum communication. *Nat. Photonics* **1**, 165–171 (2007).

- [25] Giovannetti, V., Lloyd, S. & Maccone, L. Quantum metrology. *Phys. Rev. Lett.* **96**, 010401 (2006).
- [26] Giovannetti, V., Lloyd, S. & Maccone, L. Advances in quantum metrology. *Nat. Photonics* **5**, 222–229 (2011).
- [27] Tóth, G. & Apellaniz, I. Quantum metrology from a quantum information science perspective. *J. Phys. A Math. Theor.* **47**, 424006 (2014).
- [28] Plenio, M. B. & Huelga, S. F. Dephasing-assisted transport: quantum networks and biomolecules. *New J. Phys.* **10**, 113019 (2008).
- [29] Lloyd, S. Quantum coherence in biological systems. *J. Phys. Conf. Ser.* **302**, 012037 (2011).
- [30] Huelga, S. & Plenio, M. Vibrations, quanta and biology. *Contemp. Phys.* **54**, 181–207 (2013).
- [31] Engel, G. S. *et al.* Evidence for wavelike energy transfer through quantum coherence in photosynthetic systems. *Nature* **446**, 782–786 (2007).
- [32] Rodgers, C. T. & Hore, P. J. Chemical magnetoreception in birds: The radical pair mechanism. *Proc. Natl. Acad. Sci.* **106**, 353–360 (2009).
- [33] Horodecki, R., Horodecki, P., Horodecki, M. & Horodecki, K. Quantum entanglement. *Rev. Mod. Phys.* **81**, 865–942 (2009).
- [34] Jozsa, R. & Linden, N. On the role of entanglement in quantum-computational speed-up. *Proc. R. Soc. A Math. Phys. Eng. Sci.* **459**, 2011–2032 (2003).
- [35] Bennett, C. H. & Wiesner, S. J. Communication via one- and two-particle operators on Einstein-Podolsky-Rosen states. *Phys. Rev. Lett.* **69**, 2881–2884 (1992).
- [36] Bennett, C. H. *et al.* Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels. *Phys. Rev. Lett.* **70**, 1895–1899 (1993).
- [37] Vedral, V., Plenio, M. B., Rippin, M. A. & Knight, P. L. Quantifying Entanglement. *Phys. Rev. Lett.* **78**, 2275–2279 (1997).

- [38] Vedral, V. & Plenio, M. B. Entanglement measures and purification procedures. *Phys. Rev. A* **57**, 1619–1633 (1998).
- [39] Brandão, F. G. S. L. & Plenio, M. B. Entanglement theory and the second law of thermodynamics. *Nat. Phys.* **4**, 873–877 (2008).
- [40] Baumgratz, T., Cramer, M. & Plenio, M. B. Quantifying coherence. *Phys. Rev. Lett.* **113**, 140401 (2014).
- [41] Marvian, I. & Spekkens, R. W. How to quantify coherence: Distinguishing speakable and unspeakable notions. *Phys. Rev. A* **94**, 052324 (2016).
- [42] Ollivier, H. & Zurek, W. H. Quantum Discord: A Measure of the Quantumness of Correlations. *Phys. Rev. Lett.* **88**, 017901 (2001).
- [43] Henderson, L. & Vedral, V. Classical, quantum and total correlations. *J. Phys. A: Math. Gen.* **34**, 6899–6905 (2001).
- [44] Modi, K., Brodutch, A., Cable, H., Paterek, T. & Vedral, V. The classical-quantum boundary for correlations: Discord and related measures. *Rev. Mod. Phys.* **84**, 1655–1707 (2012).
- [45] Datta, A., Shaji, A. & Caves, C. M. Quantum Discord and the Power of One Qubit. *Phys. Rev. Lett.* **100**, 050502 (2008).
- [46] Schrödinger, E. Die gegenwärtige Situation in der Quantenmechanik. *Naturwissenschaften* **23**, 807–812 (1935).
- [47] Zurek, W. H. Decoherence, einselection, and the quantum origins of the classical. *Rev. Mod. Phys.* **75**, 715–775 (2003).
- [48] Schlosshauer, M. Decoherence, the measurement problem, and interpretations of quantum mechanics. *Rev. Mod. Phys.* **76**, 1267–1305 (2005).
- [49] Kofler, J. & Brukner, Č. Classical World Arising out of Quantum Physics under the Restriction of Coarse-Grained Measurements. *Phys. Rev. Lett.* **99**, 180403 (2007).

- [50] Raeisi, S., Sekatski, P. & Simon, C. Coarse Graining Makes It Hard to See Micro-Macro Entanglement. *Phys. Rev. Lett.* **107**, 250401 (2011).
- [51] Sekatski, P., Gisin, N. & Sangouard, N. How difficult is it to prove the quantumness of macroscopic states? *Phys. Rev. Lett.* **113**, 090403 (2014).
- [52] Ghirardi, G. C., Rimini, A. & Weber, T. Unified dynamics for microscopic and macroscopic systems. *Phys. Rev. D* **34**, 470–491 (1986).
- [53] Bassi, A., Lochan, K., Satin, S., Singh, T. P. & Ulbricht, H. Models of wave-function collapse, underlying theories, and experimental tests. *Rev. Mod. Phys.* **85**, 471–527 (2013).
- [54] Nimmrichter, S. & Hornberger, K. Macroscopicity of Mechanical Quantum Superposition States. *Phys. Rev. Lett.* **110**, 160403 (2013).
- [55] Leggett, A. J. Testing the limits of quantum mechanics: motivation, state of play, prospects. *J. Phys. Condens. Matter* **14**, R415–R451 (2002).
- [56] Arndt, M. & Hornberger, K. Testing the limits of quantum mechanical superpositions. *Nat. Phys.* **10**, 271–277 (2014).
- [57] Farrow, T. & Vedral, V. Classification of macroscopic quantum effects. *Opt. Commun.* **337**, 22–26 (2015).
- [58] Arndt, M. *et al.* Wave–particle duality of C60 molecules. *Nature* **401**, 680–682 (1999).
- [59] Clarke, J. & Wilhelm, F. K. Superconducting quantum bits. *Nature* **453**, 1031–1042 (2008).
- [60] van der Wal, C. H. Quantum Superposition of Macroscopic Persistent-Current States. *Science (80-.)*. **290**, 773–777 (2000).
- [61] Friedman, J. R., Patel, V., Chen, W., Tolpygo, S. K. & Lukens, J. E. Quantum superposition of distinct macroscopic states. *Nature* **406**, 43–46 (2000).
- [62] Gao, W.-B. *et al.* Experimental demonstration of a hyper-entangled ten-qubit Schrödinger cat state. *Nat. Phys.* **6**, 331–335 (2010).

- [63] Wang, C. *et al.* A Schrodinger cat living in two boxes. *Science (80-.)*. **352**, 1087–1091 (2016).
- [64] Aspelmeyer, M., Kippenberg, T. J. & Marquardt, F. Cavity optomechanics. *Rev. Mod. Phys.* **86**, 1391–1452 (2014).
- [65] Leggett, A. J. Macroscopic Quantum Systems and the Quantum Theory of Measurement. *Suppl. Prog. Theor. Phys.* **69**, 80–100 (1980).
- [66] Björk, G. & Mana, P. G. L. A size criterion for macroscopic superposition states. *J. Opt. B Quantum Semiclass. Opt.* **6**, 429–436 (2004).
- [67] Shimizu, A. & Morimae, T. Detection of Macroscopic Entanglement by Correlation of Local Observables. *Phys. Rev. Lett.* **95**, 090401 (2005).
- [68] Marquardt, F., Abel, B. & Von Delft, J. Measuring the size of a quantum superposition of many-body states. *Phys. Rev. A* **78**, 012109 (2008).
- [69] Fröwis, F. & Dür, W. Measures of macroscopicity for quantum spin systems. *New J. Phys.* **14**, 093039 (2012).
- [70] Lee, C. W. & Jeong, H. Quantification of macroscopic quantum superpositions within phase space. *Phys. Rev. Lett.* **106**, 220401 (2011).
- [71] Dirac, P. A. M. A new notation for quantum mechanics. *Math. Proc. Cambridge Philos. Soc.* **35**, 416 (1939).
- [72] Nielsen, M. A. & Chuang, I. L. *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2010).
- [73] Wilde, M. M. *Quantum Information Theory* (Cambridge University Press, Cambridge, 2013).
- [74] Binney, J. & Skinner, D. *The Physics of Quantum Mechanics* (Oxford University Press, 2013).
- [75] Robertson, H. P. The Uncertainty Principle. *Phys. Rev.* **34**, 163–164 (1929).
- [76] Gerry, C. & Knight, P. *Introductory Quantum Optics* (Cambridge University Press, Cambridge, 2004).

- [77] Kraus, K., Böhm, A., Dollard, J. D. & Wootters, W. H. (eds.) *States, Effects, and Operations Fundamental Notions of Quantum Theory*, vol. 190 of *Lecture Notes in Physics* (Springer Berlin Heidelberg, Berlin, Heidelberg, 1983).
- [78] Stinespring, W. F. Positive functions on C^* -algebras. *Proc. Am. Math. Soc.* **6**, 211–211 (1955).
- [79] Choi, M.-D. Completely positive linear maps on complex matrices. *Linear Algebra Appl.* **10**, 285–290 (1975).
- [80] Davies, E. B. & Lewis, J. T. An operational approach to quantum probability. *Commun. Math. Phys.* **17**, 239–260 (1970).
- [81] Plenio, M. B. Logarithmic Negativity: A Full Entanglement Monotone That is not Convex. *Phys. Rev. Lett.* **95**, 090503 (2005).
- [82] Lindblad, G. On the generators of quantum dynamical semigroups. *Commun. Math. Phys.* **48**, 119–130 (1976).
- [83] Breuer, H.-P. & Petruccione, F. *The Theory of Open Quantum Systems* (Oxford University Press, 2007).
- [84] Shannon, C. E. A Mathematical Theory of Communication. *Bell Syst. Tech. J.* **27**, 623–656 (1948).
- [85] Blundell, S. J. & Blundell, K. M. *Concepts in Thermal Physics* (Oxford University Press, 2009).
- [86] Cover, T. M. & Thomas, J. A. *Elements of Information Theory* (John Wiley & Sons, Inc., Hoboken, NJ, USA, 2005).
- [87] Schumacher, B. Quantum coding. *Phys. Rev. A* **51**, 2738–2747 (1995).
- [88] Fuchs, C. A. Distinguishability and Accessible Information in Quantum Theory. *arXiv:quant-ph/9601020* (1996).
- [89] Hayashi, M. Asymptotics of quantum relative entropy from a representation theoretical viewpoint. *J. Phys. A: Math. Gen.* **34**, 3413–3419 (2001).

- [90] Braunstein, S. L. & Caves, C. M. Statistical distance and the geometry of quantum states. *Phys. Rev. Lett.* **72**, 3439–3443 (1994).
- [91] Coecke, B., Fritz, T. & Spekkens, R. W. A mathematical theory of resources. *Inf. Comput.* **250**, 59–86 (2016).
- [92] Brandão, F. G. S. L. & Gour, G. Reversible Framework for Quantum Resource Theories. *Phys. Rev. Lett.* **115**, 070503 (2015).
- [93] Horodecki, M. & Oppenheim, J. (Quantumness in the context of) Resource Theories. *Int. J. Mod. Phys. B* **27**, 1345019 (2012).
- [94] Horodecki, M. & Oppenheim, J. Fundamental limitations for quantum and nanoscale thermodynamics. *Nat. Commun.* **4**, 2059 (2013).
- [95] Gallego, R. & Aolita, L. Resource Theory of Steering. *Phys. Rev. X* **5**, 041008 (2015).
- [96] Horodecki, M., Shor, P. W. & Ruskai, M. B. Entanglement Breaking Channels. *Rev. Math. Phys.* **15**, 629–641 (2003).
- [97] Chitambar, E., Leung, D., Mančinska, L., Ozols, M. & Winter, A. Everything You Always Wanted to Know About LOCC (But Were Afraid to Ask). *Commun. Math. Phys.* **328**, 303–326 (2014).
- [98] Rains, E. Entanglement purification via separable superoperators. *arXiv:quant-ph/9707002v3* (1997).
- [99] Nielsen, M. A. Conditions for a Class of Entanglement Transformations. *Phys. Rev. Lett.* **83**, 436–439 (1999).
- [100] Dür, W., Vidal, G. & Cirac, J. I. Three qubits can be entangled in two inequivalent ways. *Phys. Rev. A* **62**, 062314 (2000).
- [101] Gour, G. & Spekkens, R. W. The resource theory of quantum reference frames: manipulations and monotones. *New J. Phys.* **10**, 033023 (2008).
- [102] Leggett, A. J. & Garg, A. Quantum mechanics versus macroscopic realism: Is the flux there when nobody looks? *Phys. Rev. Lett.* **54**, 857–860 (1985).

- [103] Emary, C., Lambert, N. & Nori, F. Leggett-Garg Inequalities. *Reports Prog. Phys.* **77**, 016001 (2013).
- [104] Horodecki, M., Oppenheim, J. & Horodecki, R. Are the Laws of Entanglement Theory Thermodynamical? *Phys. Rev. Lett.* **89**, 240403 (2002).
- [105] Wilde, M. M. & Mizel, A. Addressing the Clumsiness Loophole in a Leggett-Garg Test of Macrorealism. *Found. Phys.* **42**, 256–265 (2012).
- [106] Maroney, O. J. E. & Timpson, C. G. Quantum- vs. Macro- Realism: What does the Leggett-Garg Inequality actually test? *arXiv:1412.6139* (2014).
- [107] Knee, G. C. *et al.* A strict experimental test of macroscopic realism in a superconducting flux qubit. *Nat. Commun.* **7**, 13253 (2016).
- [108] Benatti, F., Ghirardi, G. C., Rimini, A. & Weber, T. Operations involving momentum variables in non-Hamiltonian evolution equations. *Nuovo Cim. B* **101**, 333–355 (1988).
- [109] Milburn, G. J. Intrinsic decoherence in quantum mechanics. *Phys. Rev. A* **44**, 5401–5406 (1991).
- [110] Greenberger, D. M., Horne, M. A. & Zeilinger, A. Going Beyond Bell’s Theorem. In *Bell’s Theorem, Quantum Theory and Conceptions of the Universe*, 69–72 (Springer Netherlands, Dordrecht, 1989).
- [111] Korsbakken, J. I., Whaley, K. B., Dubois, J. & Cirac, J. I. Measurement-based measure of the size of macroscopic quantum superpositions. *Phys. Rev. A* **75**, 042106 (2007).
- [112] Sekatski, P., Sangouard, N. & Gisin, N. Size of quantum superpositions as measured with classical detectors. *Phys. Rev. A - At. Mol. Opt. Phys.* **89**, 012116 (2014).
- [113] Jeong, H., Noh, C., Bae, S., Angelakis, D. G. & Ralph, T. C. Detecting the degree of macroscopic quantumness using an overlap measurement. *J. Opt. Soc. Am. B* **31**, 3057 (2014).

- [114] Peres, A. & Scudo, P. F. Unspeakable quantum information. *Quantum Theory Reconsideration Found.* **2**, 283–298 (2002).
- [115] Bartlett, S. D., Rudolph, T. & Spekkens, R. W. Reference frames, superselection rules, and quantum information. *Rev. Mod. Phys.* **79**, 555–609 (2007).
- [116] Åberg, J. Quantifying Superposition. *arXiv:quant-ph/0612146* (2006).
- [117] Winter, A. & Yang, D. Operational Resource Theory of Coherence. *Phys. Rev. Lett.* **116**, 120404 (2016).
- [118] Chitambar, E. *et al.* Assisted Distillation of Quantum Coherence. *Phys. Rev. Lett.* **116**, 1–10 (2016).
- [119] Streltsov, A., Rana, S., Bera, M. N. & Lewenstein, M. Towards Resource Theory of Coherence in Distributed Scenarios. *Phys. Rev. X* **7**, 011024 (2017).
- [120] Streltsov, A. *et al.* Entanglement and Coherence in Quantum State Merging. *Phys. Rev. Lett.* **116**, 240405 (2016).
- [121] Bai, Z. & Du, S. Maximally coherent states. *arXiv:1503.07103* (2015).
- [122] Peng, Y., Jiang, Y. & Fan, H. Maximally coherent states and coherence-preserving operations. *Phys. Rev. A* **93**, 032326 (2016).
- [123] Napoli, C. *et al.* Robustness of Coherence: An Operational and Observable Measure of Quantum Coherence. *Phys. Rev. Lett.* **116**, 150502 (2016).
- [124] Vatasescu, M. Measures of electronic-vibrational entanglement and quantum coherence in a molecular system. *Phys. Rev. A* **92**, 042323 (2015).
- [125] Roden, J. J. J. & Whaley, K. B. Probability-current analysis of energy transport in open quantum systems. *Phys. Rev. E* **93**, 012128 (2016).
- [126] Karpat, G., Çakmak, B. & Fanchini, F. F. Quantum coherence and uncertainty in the anisotropic XY chain. *Phys. Rev. B* **90**, 104431 (2014).
- [127] Malvezzi, A. L. *et al.* Quantum correlations and coherence in spin-1 Heisenberg chains. *Phys. Rev. B* **93**, 184428 (2016).

- [128] Li, Y.-C. & Lin, H.-Q. Quantum coherence and quantum phase transitions. *Sci. Rep.* **6**, 26365 (2016).
- [129] Streltsov, A. Genuine Quantum Coherence. *arXiv:1511.08346* (2015).
- [130] Chitambar, E. & Gour, G. Critical Examination of Incoherent Operations and a Physically Consistent Resource Theory of Quantum Coherence. *Phys. Rev. Lett.* **117**, 030401 (2016).
- [131] de Vicente, J. I. & Streltsov, A. The power of the resource theory of genuine quantum coherence. *arXiv:1604.08031* (2016).
- [132] Hu, X. Channels that do not generate coherence. *Phys. Rev. A* **94**, 012326 (2016).
- [133] Levi, F. & Mintert, F. A quantitative theory of coherent delocalization. *New J. Phys.* **16**, 033007 (2014).
- [134] Brandão, F. G. S. L., Horodecki, M., Oppenheim, J., Renes, J. M. & Spekkens, R. W. Resource Theory of Quantum States Out of Thermal Equilibrium. *Phys. Rev. Lett.* **111**, 250404 (2013).
- [135] Lostaglio, M., Jennings, D. & Rudolph, T. Description of quantum coherence in thermodynamic processes requires constraints beyond free energy. *Nat. Commun.* **6**, 6383 (2015).
- [136] Horn, R. A. & Johnson, C. R. *Topics in matrix analysis*, vol. 35 (Cambridge University Press, Cambridge, 1991).
- [137] Shao, L.-H., Xi, Z., Fan, H. & Li, Y. Fidelity and trace-norm distances for quantifying coherence. *Phys. Rev. A* **91**, 042120 (2015).
- [138] Streltsov, A., Singh, U., Dhar, H. S., Bera, M. N. & Adesso, G. Measuring Quantum Coherence with Entanglement. *Phys. Rev. Lett.* **115**, 020403 (2015).
- [139] Rana, S., Parashar, P. & Lewenstein, M. Trace-distance measure of coherence. *Phys. Rev. A* **93**, 012110 (2016).

- [140] Keyl, M. & Werner, R. F. Optimal cloning of pure states, testing single clones. *J. Math. Phys.* **40**, 3283 (1999).
- [141] Marvian, I. & Mann, R. B. Building all time evolutions with rotationally invariant Hamiltonians. *Phys. Rev. A* **78**, 022304 (2008).
- [142] Ahmadi, M., Jennings, D. & Rudolph, T. The Wigner–Araki–Yanase theorem and the quantum resource theory of asymmetry. *New J. Phys.* **15**, 013057 (2013).
- [143] Navascués, M. & Popescu, S. How Energy Conservation Limits Our Measurements. *Phys. Rev. Lett.* **112**, 140502 (2014).
- [144] Åberg, J. Catalytic Coherence. *Phys. Rev. Lett.* **113**, 150402 (2014).
- [145] Brandão, F., Horodecki, M., Ng, N., Oppenheim, J. & Wehner, S. The second laws of quantum thermodynamics. *Proc. Natl. Acad. Sci.* **112**, 3275–3279 (2015).
- [146] Lostaglio, M., Korzekwa, K., Jennings, D. & Rudolph, T. Quantum Coherence, Time-Translation Symmetry, and Thermodynamics. *Phys. Rev. X* **5**, 021001 (2015).
- [147] Korzekwa, K., Lostaglio, M., Oppenheim, J. & Jennings, D. The extraction of work from quantum coherence. *New J. Phys.* **18**, 023045 (2016).
- [148] Holevo, A. S. *Quantum Systems, Channels, Information* (de Gruyter, Berlin, Boston, 2012).
- [149] Narasimhachar, V. & Gour, G. Low-temperature thermodynamics with quantum coherence. *Nat. Commun.* **6**, 7689 (2015).
- [150] Chitambar, E. & Hsieh, M.-H. Relating the Resource Theories of Entanglement and Quantum Coherence. *Phys. Rev. Lett.* **117**, 020402 (2016).
- [151] Adesso, G., Bromley, T. R. & Cianciaruso, M. Measures and applications of quantum correlations. *arXiv:1605.00806* (2016).

- [152] Ferraro, A., Aolita, L., Cavalcanti, D., Cucchietti, F. M. & Acín, A. Almost all quantum states have nonclassical correlations. *Phys. Rev. A* **81**, 052318 (2010).
- [153] Yu, T. & Eberly, J. H. Sudden Death of Entanglement. *Science (80-.)*. **323**, 598–601 (2009).
- [154] Madhok, V. & Datta, A. Interpreting quantum discord through quantum state merging. *Phys. Rev. A* **83**, 032323 (2011).
- [155] Cavalcanti, D. *et al.* Operational interpretations of quantum discord. *Phys. Rev. A* **83**, 032324 (2011).
- [156] Dakić, B. *et al.* Quantum discord as resource for remote state preparation. *Nat. Phys.* **8**, 666–670 (2012).
- [157] Pirandola, S. Quantum discord as a resource for quantum cryptography. *Sci. Rep.* **4**, 6956 (2014).
- [158] Gu, M. *et al.* Observing the operational significance of discord consumption. *Nat. Phys.* **8**, 671–675 (2012).
- [159] Weedbrook, C., Pirandola, S., Thompson, J., Vedral, V. & Gu, M. How discord underlies the noise resilience of quantum illumination. *New J. Phys.* **18**, 043027 (2016).
- [160] Boixo, S., Aolita, L., Cavalcanti, D., Modi, K. & Winter, A. Quantum Locking of Classical Correlations and Quantum Discord Classical-Quantum States. *Int. J. Quantum Inf.* **09**, 1643–1651 (2011).
- [161] Oppenheim, J., Horodecki, M., Horodecki, P. & Horodecki, R. Thermodynamical Approach to Quantifying Quantum Correlations. *Phys. Rev. Lett.* **89**, 180402 (2002).
- [162] Horodecki, M. *et al.* Local versus nonlocal information in quantum-information theory: Formalism and phenomena. *Phys. Rev. A* **71**, 062307 (2005).
- [163] Knill, E. & Laflamme, R. Power of One Bit of Quantum Information. *Phys. Rev. Lett.* **81**, 5672–5675 (1998).

- [164] Streltsov, A., Kampermann, H. & Bruß, D. Behavior of Quantum Correlations under Local Noise. *Phys. Rev. Lett.* **107**, 170502 (2011).
- [165] Vedral, V. The role of relative entropy in quantum information theory. *Rev. Mod. Phys.* **74**, 197–234 (2002).
- [166] Datta, A. *Studies on the Role of Entanglement in Mixed-state Quantum Computation*. Ph.D. thesis, arXiv:0807.4490 (2008).
- [167] Wu, S., Poulsen, U. V. & Mølmer, K. Correlations in local measurements on a quantum state, and complementarity as an explanation of nonclassicality. *Phys. Rev. A* **80**, 032319 (2009).
- [168] Lang, M. D., Caves, C. M. & Shaji, A. Entropic measures of non-classical correlations. *Int. J. Quantum Inf.* **09**, 1553–1586 (2011).
- [169] Piani, M., Horodecki, P. & Horodecki, R. No-Local-Broadcasting Theorem for Multipartite Quantum Correlations. *Phys. Rev. Lett.* **100**, 090502 (2008).
- [170] Piani, M. & Adesso, G. Quantumness of correlations revealed in local measurements exceeds entanglement. *Phys. Rev. A* **85**, 040301 (2012).
- [171] Yao, Y., Xiao, X., Ge, L. & Sun, C. P. Quantum coherence in multipartite systems. *Phys. Rev. A* **92**, 022112 (2015).
- [172] Modi, K., Paterek, T., Son, W., Vedral, V. & Williamson, M. Unified View of Quantum and Classical Correlations. *Phys. Rev. Lett.* **104**, 080501 (2010).
- [173] Piani, M. *et al.* All Nonclassical Correlations Can Be Activated into Distillable Entanglement. *Phys. Rev. Lett.* **106**, 220403 (2011).
- [174] Rulli, C. C. & Sarandy, M. S. Global quantum discord in multipartite systems. *Phys. Rev. A* **84**, 042109 (2011).
- [175] Datta, A. & Vidal, G. Role of entanglement and correlations in mixed-state quantum computation. *Phys. Rev. A* **75**, 042310 (2007).
- [176] Datta, A., Flammia, S. T. & Caves, C. M. Entanglement and the power of one qubit. *Phys. Rev. A* **72**, 042316 (2005).

- [177] Dakić, B., Vedral, V. & Brukner, Č. Necessary and Sufficient Condition for Nonzero Quantum Discord. *Phys. Rev. Lett.* **105**, 190502 (2010).
- [178] Yu, S., Zhang, C., Chen, Q. & Oh, C. H. Quantum channels that preserve the commutativity. *arXiv:1112.5700* (2011).
- [179] Hu, X., Fan, H., Zhou, D. L. & Liu, W.-M. Necessary and sufficient conditions for local creation of quantum correlation. *Phys. Rev. A* **85**, 032102 (2012).
- [180] Guo, Y. & Hou, J. Necessary and sufficient conditions for the local creation of quantum discord. *J. Phys. A Math. Theor.* **46**, 155301 (2013).
- [181] Petz, D. Monotonicity of quantum relative entropy revisited. *Rev. Math. Phys.* **15**, 79–91 (2003).
- [182] Piani, M. Problem with geometric discord. *Phys. Rev. A* **86**, 034101 (2012).
- [183] Seshadreesan, K. P. & Wilde, M. M. Fidelity of recovery, squashed entanglement, and measurement recoverability. *Phys. Rev. A* **92**, 042321 (2015).
- [184] Matera, J. M., Egloff, D., Killoran, N. & Plenio, M. B. Coherent control of quantum systems as a resource theory. *Quantum Sci. Technol.* **1**, 01LT01 (2016).
- [185] Marvian, I. & Spekkens, R. W. Asymmetry properties of pure quantum states. *Phys. Rev. A* **90**, 014102 (2014).
- [186] Marvian, I. & Spekkens, R. W. Extending Noether’s theorem by quantifying the asymmetry of quantum states. *Nat. Commun.* **5**, 6 (2014).
- [187] Marvian, I. & Spekkens, R. W. Modes of asymmetry: The application of harmonic analysis to symmetric quantum dynamics and quantum reference frames. *Phys. Rev. A* **90**, 062110 (2014).
- [188] Bartlett, S. D., Rudolph, T. & Spekkens, R. W. Classical and Quantum Communication without a Shared Reference Frame. *Phys. Rev. Lett.* **91**, 027901 (2003).

- [189] Ball, J. L. & Banaszek, K. Hybrid noiseless subsystems for quantum communication over optical fibres. *J. Phys. A. Math. Gen.* **39**, L1–L7 (2006).
- [190] Javanainen, J. & Yoo, S. M. Quantum Phase of a Bose-Einstein Condensate with an Arbitrary Number of Atoms. *Phys. Rev. Lett.* **76**, 161–164 (1996).
- [191] Barnett, S., Burnett, K. & Vaccaro, J. Why a condensate can be thought of as having a definite phase. *J. Res. Natl. Inst. Stand. Technol.* **101**, 593 (1996).
- [192] Bartlett, S. D., Rudolph, T. & Spekkens, R. W. Dialogue concerning two views on quantum coherence: factist and fictionist. *Int. J. Quantum Inf.* **04**, 17–43 (2006).
- [193] Dowling, M. R., Bartlett, S. D., Rudolph, T. & Spekkens, R. W. Observing a coherent superposition of an atom and a molecule. *Phys. Rev. A* **74**, 052113 (2006).
- [194] Gour, G., Marvian, I. & Spekkens, R. W. Measuring the quality of a quantum reference frame: The relative entropy of frameness. *Phys. Rev. A* **80**, 012307 (2009).
- [195] Tóth, G. & Petz, D. Extremal properties of the variance and the quantum Fisher information. *Phys. Rev. A* **87**, 032324 (2013).
- [196] Yu, S. Quantum Fisher Information as the Convex Roof of Variance. *arXiv:1302.5311* (2013).
- [197] Morozova, E. A. & Chentsov, N. N. Markov invariant geometry on manifolds of states. *J. Sov. Math.* **56**, 2648–2669 (1991).
- [198] Petz, D. Covariance and Fisher information in quantum mechanics. *J. Phys. A. Math. Gen.* **35**, 929–939 (2002).
- [199] Petz, D. Monotone metrics on matrix spaces. *Linear Algebra Appl.* **244**, 81–96 (1996).
- [200] Wigner, E. P. & Yanase, M. M. Information contents of distributions. *Proc. Natl. Acad. Sci. U. S. A.* **49**, 910 (1963).

- [201] Girolami, D. Observable Measure of Quantum Coherence in Finite Dimensional Systems. *Phys. Rev. Lett.* **113**, 170401 (2014).
- [202] Gibilisco, P., Imparato, D. & Isola, T. Inequalities for quantum Fisher information. *Proc. Am. Math. Soc.* **137**, 317–327 (2008).
- [203] Schuch, N., Verstraete, F. & Cirac, J. I. Quantum entanglement theory in the presence of superselection rules. *Phys. Rev. A* **70**, 042310 (2004).
- [204] Petrov, V. V. *Sums of Independent Random Variables* (Springer Berlin Heidelberg, Berlin, Heidelberg, 1975).
- [205] Wootters, W. K. Entanglement of Formation and Concurrence. *Quantum Info. Comput.* **1**, 27–44 (2001).
- [206] Toloui, B., Gour, G. & Sanders, B. C. Constructing monotones for quantum phase references in totally dephasing channels. *Phys. Rev. A - At. Mol. Opt. Phys.* **84**, 1–8 (2011).
- [207] Morimae, T., Sugita, A. & Shimizu, A. Macroscopic entanglement of many-magnon states. *Phys. Rev. A* **71**, 032317 (2005).
- [208] Morimae, T. Superposition of macroscopically distinct states means large multipartite entanglement. *Phys. Rev. A* **81**, 010101 (2010).
- [209] Fröwis, F., Sangouard, N. & Gisin, N. Linking measures for macroscopic quantum states via photon–spin mapping. *Opt. Commun.* **337**, 2–11 (2015).
- [210] Oudot, E., Sekatski, P., Fröwis, F., Gisin, N. & Sangouard, N. Two-mode squeezed states as Schrödinger cat-like states. *J. Opt. Soc. Am. B* **32**, 2190 (2015).
- [211] Stoler, D. Equivalence Classes of Minimum Uncertainty Packets. *Phys. Rev. D* **1**, 3217–3219 (1970).
- [212] Fröwis, F., Schmied, R. & Gisin, N. Tighter quantum uncertainty relations following from a general probabilistic bound. *Phys. Rev. A* **92**, 012102 (2015).

- [213] Jeong, H., Kang, M. & Lee, C.-W. Reply to Comment on “Quantification of Macroscopic Quantum Superpositions within Phase Space”. *arXiv:1108.0212* (2011).
- [214] Ozawa, M. Entanglement measures and the Hilbert–Schmidt distance. *Phys. Lett. A* **268**, 158–160 (2000).
- [215] Nimmrichter, S. *Macroscopic Matter Wave Interferometry*. Springer Theses (Springer International Publishing, Cham, 2014).
- [216] Tiranov, A. *et al.* Demonstration of Light-Matter Micro-Macro Quantum Correlations. *Phys. Rev. Lett.* **116**, 190502 (2016).
- [217] Stam, A. Some inequalities satisfied by the quantities of information of Fisher and Shannon. *Inf. Control* **2**, 101–112 (1959).
- [218] Sutter, D., Tomamichel, M. & Harrow, A. W. Strengthened Monotonicity of Relative Entropy via Pinched Petz Recovery Map. *IEEE Trans. Inf. Theory* **62**, 2907–2913 (2016).
- [219] Petz, D. & Ghinea, C. Introduction to quantum Fisher information. *arXiv:1008.2417* (2010).
- [220] König, R. & Smith, G. The Entropy Power Inequality for Quantum Systems. *IEEE Trans. Inf. Theory* **60**, 1536–1548 (2014).
- [221] Carlson, B. C. Some inequalities for hypergeometric functions. *Proc. Am. Math. Soc.* **17**, 32–32 (1966).
- [222] Kwon, H., Park, C.-Y., Tan, K. C. & Jeong, H. Disturbance-Based Measure of Macroscopic Coherence. *arXiv:1608.01122* (2016).
- [223] Fröwis, F. Lower bounds on the size of general Schrödinger-cat states from experimental data. *J. Phys. A Math. Theor.* **50**, 114003 (2017).
- [224] Gühne, O., Tóth, G. & Briegel, H. J. Multipartite entanglement in spin chains. *New J. Phys.* **7**, 229–229 (2005).

- [225] Tóth, G. Multipartite entanglement and high-precision metrology. *Phys. Rev. A* **85**, 022322 (2012).
- [226] Hyllus, P. *et al.* Fisher information and multiparticle entanglement. *Phys. Rev. A* **85**, 022321 (2012).
- [227] Pezzé, L. & Smerzi, A. Entanglement, Nonlinear Dynamics, and the Heisenberg Limit. *Phys. Rev. Lett.* **102**, 100401 (2009).
- [228] Apellaniz, I., Kleinmann, M., Gühne, O. & Tóth, G. Optimal witnessing of the quantum Fisher information with few measurements. *Phys. Rev. A* **95**, 032330 (2017).
- [229] Gessner, M., Pezzè, L. & Smerzi, A. Resolution-enhanced entanglement detection. *Phys. Rev. A* **95**, 032326 (2017).
- [230] Augusiak, R. *et al.* Asymptotic role of entanglement in quantum metrology. *Phys. Rev. A* **94**, 012339 (2016).
- [231] Leggett, A. J. Note on the "size" of Schroedinger cats. *arXiv:1603.03992* (2016).
- [232] Korsbakken, J. I., Wilhelm, F. K. & Whaley, K. B. The size of macroscopic superposition states in flux qubits. *EPL (Europhysics Lett.)* **89**, 30003 (2010).
- [233] Cartz, L. Thermal Vibrations of Atoms in Cubic Crystals II: The Amplitude of Atomic Vibrations. *Proc. Phys. Soc. Sect. B* **68**, 957–967 (1955).
- [234] Dür, W., Simon, C. & Cirac, J. I. Effective Size of Certain Macroscopic Quantum Superpositions. *Phys. Rev. Lett.* **89**, 210402 (2002).
- [235] Hornberger, K., Gerlich, S., Haslinger, P., Nimmrichter, S. & Arndt, M. Colloquium : Quantum interference of clusters and molecules. *Rev. Mod. Phys.* **84**, 157–173 (2012).
- [236] Juffmann, T. *et al.* Real-time single-molecule imaging of quantum interference. *Nat. Nanotechnol.* **7**, 297–300 (2012).

- [237] Eibenberger, S., Gerlich, S., Arndt, M., Mayor, M. & Tüxen, J. Matter-wave interference of particles selected from a molecular library with masses exceeding 10 000 amu. *Phys. Chem. Chem. Phys.* **15**, 14696 (2013).
- [238] Zeilinger, A., Gaehler, R., Shull, C. G. & Treimer, W. Experimental status and recent results of neutron interference optics. *AIP Conf. Proc.* **89**, 93–100 (1982).
- [239] Ringbauer, M., Weinhold, T. J., White, A. G. & Vanner, M. R. Generation of Mechanical Interference Fringes by Multi-Photon Quantum Measurement. *arXiv:1602.05955* (2016).
- [240] Marshall, W., Simon, C., Penrose, R. & Bouwmeester, D. Towards Quantum Superpositions of a Mirror. *Phys. Rev. Lett.* **91**, 130401 (2003).
- [241] Pikovski, I., Vanner, M. R., Aspelmeyer, M., Kim, M. S. & Brukner, Č. Probing Planck-scale physics with quantum optics. *Nat. Phys.* **8**, 393–397 (2012).
- [242] Teufel, J. D. *et al.* Sideband cooling of micromechanical motion to the quantum ground state. *Nature* **475**, 359–363 (2011).
- [243] Chan, J. *et al.* Laser cooling of a nanomechanical oscillator into its quantum ground state. *Nature* **478**, 89–92 (2011).
- [244] Verhagen, E., Deléglise, S., Weis, S., Schliesser, A. & Kippenberg, T. J. Quantum-coherent coupling of a mechanical oscillator to an optical cavity mode. *Nature* **482**, 63–67 (2012).
- [245] Bose, S. *et al.* A Spin Entanglement Witness for Quantum Gravity. *arXiv:1707.06050* (2017).
- [246] Korsbakken, J. I., Wilhelm, F. K. & Whaley, K. B. Electronic structure of superposition states in flux qubits. *Phys. Scr.* **T137**, 014022 (2009).
- [247] Altland, A. & Simons, B. D. *Condensed Matter Field Theory* (Cambridge University Press, Cambridge, 2010).
- [248] Vedral, V. Quantifying entanglement in macroscopic systems. *Nature* **453**, 1004–1007 (2008).

- [249] Andrews, M. R. *et al.* Observation of Interference Between Two Bose Condensates. *Science (80-.)*. **275**, 637–641 (1997).
- [250] Bravyi, S. & Kitaev, A. Universal quantum computation with ideal Clifford gates and noisy ancillas. *Phys. Rev. A* **71**, 022316 (2005).
- [251] Hayden, P. M., Horodecki, M. & Terhal, B. M. The asymptotic entanglement cost of preparing a quantum state. *J. Phys. A. Math. Gen.* **34**, 6891–6898 (2001).
- [252] Hauke, P., Heyl, M., Tagliacozzo, L. & Zoller, P. Measuring multipartite entanglement through dynamic susceptibilities. *Nat. Phys.* **12**, 778–782 (2016).
- [253] Fröwis, F., Sekatski, P. & Dür, W. Detecting Large Quantum Fisher Information with Finite Measurement Precision. *Phys. Rev. Lett.* **116**, 090801 (2016).
- [254] Killoran, N., Steinhoff, F. E. S. & Plenio, M. B. Converting Nonclassicality into Entanglement. *Phys. Rev. Lett.* **116**, 080402 (2016).