A dialogue between quantum information and thermodynamics

Submitted December 2020 (resubmitted April 2021), in fulfillment of
the conditions for the award of the degree **PhD in Mathematics.**

Benjamin Jordan Morris

4302149

Supervised by Prof. Gerardo Adesso

School of Mathematical Sciences

University of Nottingham

I hereby declare that this dissertation is all my own work, except as indicated in the
text:

Signature _____

Date _____ / _____ / _____

I hereby declare that I have all necessary rights and consents to publicly distribute this
dissertation via the University of Nottingham's e-dissertation archive.

In 1949, when Shannon had been working on his equations for some time, he happened to visit the mathematician John von Neumann, who asked him how he was getting on with his theory of missing information. Shannon replied that the theory was in excellent shape, except that he needed a good name for "missing information". "Why dont you call it entropy ", von Neumann suggested. "In the first place, a mathematical development very much like yours already exists in Boltzmanns statistical mechanics, and in the second place, no one understands entropy very well, so in any discussion you will be in a position of advantage. [15]"

Abstract

Thermodynamics has always been inextricably linked with the abstract concept of information. The past decades have proved this relationship fruitful once again with the incorporation of many concepts and ideas from quantum information theory into the context of thermodynamics. This has allowed us to better understand the role played by uniquely quantum features such as coherence and entanglement in thermodynamics and encouraged us to test some of the fundamental limitations of thermodynamics within quantum physics laboratories. In this thesis I will demonstrate some of these fruits that allow us to both reformulate and test thermodynamics while forcing us to operationally understand the thermodynamic implications of our abstract information based results.

Chapter 1 begins by following the historical and conceptual bridge linking thermodynamics and information, in particular how our usage of entropy in defining our underlying particle ensembles naturally imbues our theory of thermodynamics with information theoretic overtones. I will also introduce many of the tools and concepts that will be employed in the following chapters, including resource theoretic formulations, quantum particles and fluctuation relations.

In chapter 2 we study the process of assisted work distillation. This scenario arises when two parties share a bipartite quantum state and their task is to locally distill the optimal amount of work when one party is restricted to thermal operations, whereas the other can perform general quantum operations and they are allowed to communicate classically. We find that this question is intimately related to the distillation of classical and quantum correlations.

In chapter 3 we investigate manipulations of pure quantum states under incoherent or strictly incoherent operations assisted by a coherence battery, that is, a storage device whose degree of coherence is allowed to fluctuate in the process. This leads to the derivation of fluctuation relations for quantum coherence, analogous to Jarzynski's and Crooks' relations for work in thermodynamics.

In chapter 4 we study a quantum analogue of the famous classical Gibbs paradox. This paradox forces us to take a closer look at our notion of distinguishability and the role of the observer in classical thermodynamics. Namely will an observer calculate an entropy change when two different classical gasses mix if, for said observer, the gasses cannot be distinguished. By moving the thought experiment into the quantum realm, we reveal new and surprising behaviour. We show that the ignorant observer, who cannot distinguish the gases with devices in their lab, can in fact extract work from mixing. This effect demonstrates the importance of carefully accounting for the level of knowledge of an

observer, and its implications for genuinely quantum modifications to thermodynamics.

In the final chapter 5 we look further at the properties of these identical quantum particles. In particular, because of their exchange symmetry, identical particles can appear to be entangled–where a complete description of a physical system cannot be gained from an understanding of its parts. However, a long-standing debate has questioned whether identical particle entanglement is physical or merely a mathematical artefact. In this chapter we provide such particle entanglement with a consistent theoretical description which we believe provides the resolutive step in this enduring debate and solidify our claim by using our tools to provide the first experimental quantitative estimation of identical particle entanglement.

It is hoped that the content of these chapters will both inform and convince the reader that the role of quantum information in thermodynamics is complex and fruitful. That when we look deeper at the implicit presence of information in our theory of thermodynamics we can better understand how such a theory may be consistently merged with quantum theory. In addition, it is hoped that with the specific focus on the state space behaviour of quantum particles and the chapter studying the incongruous behaviour of a quantum Gibb's paradox that this thesis may motivate further studies into the emerging field of many-body quantum thermodynamics.

Acknowledgements

I should naturally start with my supervisor Prof. Gerardo Adesso who, despite it quickly becoming clear that I had bluffed my way through my PhD interview, never stopped encouraging me and providing me with opportunities to become a better scientist. It is also definitely true that the majority of this thesis would be incredibly lacking without the direction and input of two particularly important post-docs within the quantum correlations group, Dr. Benjamin Yadin and Dr. Ludovico Lami, who both demonstrated the right amounts of rigour and humbleness to become excellent and kind scientists. The wider support of the quantum correlations group also proved an invaluable ingredient in both my work and personal life.

I must also single out the Physical Institute for Theoretical Hierarchy (PITH) and its members who provided the original motivation and essence for many of the works in this thesis. Although not directly involved with my PhD, it was Dr. David Bruschi, my masters project supervisor, that convinced me I knew nothing of quantum mechanics and took it upon himself to change that. It would also be remiss of me not to thank the operational team at the Nottingham School of Mathematical Sciences, who got the department and my office safely up and running during the Covid-19 global pandemic. A big thank you to all my family and friends, who despite never really understanding what I was doing, encouraged me anyway. And finally, I need to acknowledge the support of Bethan Morris, who started her PhD as an office mate and finished as my wife.

List of publications

Please find below a list of publications and pre-prints by the present author [169, 168, 249, 170] which have formed the basis of chapters [2,3,4,5] sequentially.

1. <u>B. Morris</u>, L.Lami and G. Adesso, "Assisted work distillation", *Physical Review Letters* , 122(13), 130601. (2019). Ref [169]

2. <u>B. Morris</u> and G. Adesso, "Quantum coherence fluctuation relations", *Journal of Physics A: Mathematical and Theoretical* , Volume 51, Number 41 (2018). Ref [168]

3. <u>B. Morris</u>, B.Yadin and G. Adesso, "Extracting work from mixing indistinguishable systems: A quantum Gibbs "paradox" ", *arXiv preprint, currently in submission (Nature Communications)* , (2020). Ref [249]

4. <u>B. Morris</u>, B.Yadin, M.Fadel, T.Zibold, P.Treutlein and G. Adesso, "Entanglement between identical particles is a useful and consistent resource", *Physical Review X* **10**, 041012 (2020). Ref [170]

Other works by the present author which are not explicitly included in this thesis.

1. L. Correa, B. Xu, <u>B. Morris</u> and G. Adesso, "Pushing the limits of the reaction-coordinate mapping", *Journal of Chemical Physics* , 151, 094107 (2019). Ref [62]

2. D.E.Bruschi <u>B. Morris</u> and I.Fuentes, "Thermodynamics of relativistic quantum fields confined in cavities", *Physics Letters A* , 126601 (2020). Ref [48]

3. G.Landi, G.Guarnieri <u>B. Morris</u>, J.Goold and G.Adesso, "A resource theory of Maxwell's demons", *arXiv preprint* , 1st of April (2019). Ref [152]

An updated list of publications can be found via the following link

Declaration of originality

I hereby declare that the following thesis is composed of my own original work, developed during my PhD at University of Nottingham in the Quantum Correlations group. All published and unpublished work of others has been appropriately referenced and acknowledged.

Declaration of copyright

Contents

Chapter 1

Introduction

1.1 Opening remarks

When being asked by senior physicists what my PhD was on I proudly told them that I study *"Quantum Thermodynamics"*, what Prof. Adesso had failed to tell me when I agreed to conduct a PhD on the topic was that quantum thermodynamics didn't really exist as a subject in its own right and more often than not was told this by the physicists questioning me. Perhaps they had a point, what did heat and pistons and temperature have to do with qubits and entanglement and exclusion principles. In fact Einstein himself was once quoted as saying:

> *"A law is more impressive the greater the simplicity of its premises, the more different are the kinds of things it relates, and the more extended its range of applicability. (Thermodynamics) is the only physical theory of universal content, which I am convinced, that within the framework of applicability of its basic concepts will never be overthrown."*

Enter me and my four years of undergraduate physics training.

However, if those physicists were right and quantum thermodynamics was a dichotomy of terms, why was it a dichotomy of terms? What makes quantum so special as to exist outside the realm of thermodynamics? Or conversely, what makes thermodynamics so special as to work independently of quantum mechanics? The solution to these questions and the hierarchical battle underlying them is by no means an answered one, at least not in this thesis.

Nevertheless, despite the canyon separating these two behemoths of modern and clas-

sical physics we can at least observe a bridge that allows us to hop from one side to the other, *Entropy*! As responsible for measuring entanglement as it is for defining the laws of thermodynamics and seemingly as ubiquitous in quantum mechanics as it is in thermodynamics. This quantity, for many of the works discussed in this thesis, is the very bridge holding together the quantum and the thermodynamics. Unfortunately, this conceptual bridge has, in the past, been described as "The most difficult concept to understand in the whole field of physics [206] ", but entropy is meant to be a measurement of how much we don't understand something, so I probably shouldn't worry too much.

1.2 Foundational thermodynamics

Let us look more deeply at this bridge of entropy connecting thermodynamics and quantum mechanics 1.2. What have been the historical and conceptual steps along the way? Following this train of thought will motivate many of the topics that will later form the cornerstones of this thesis.
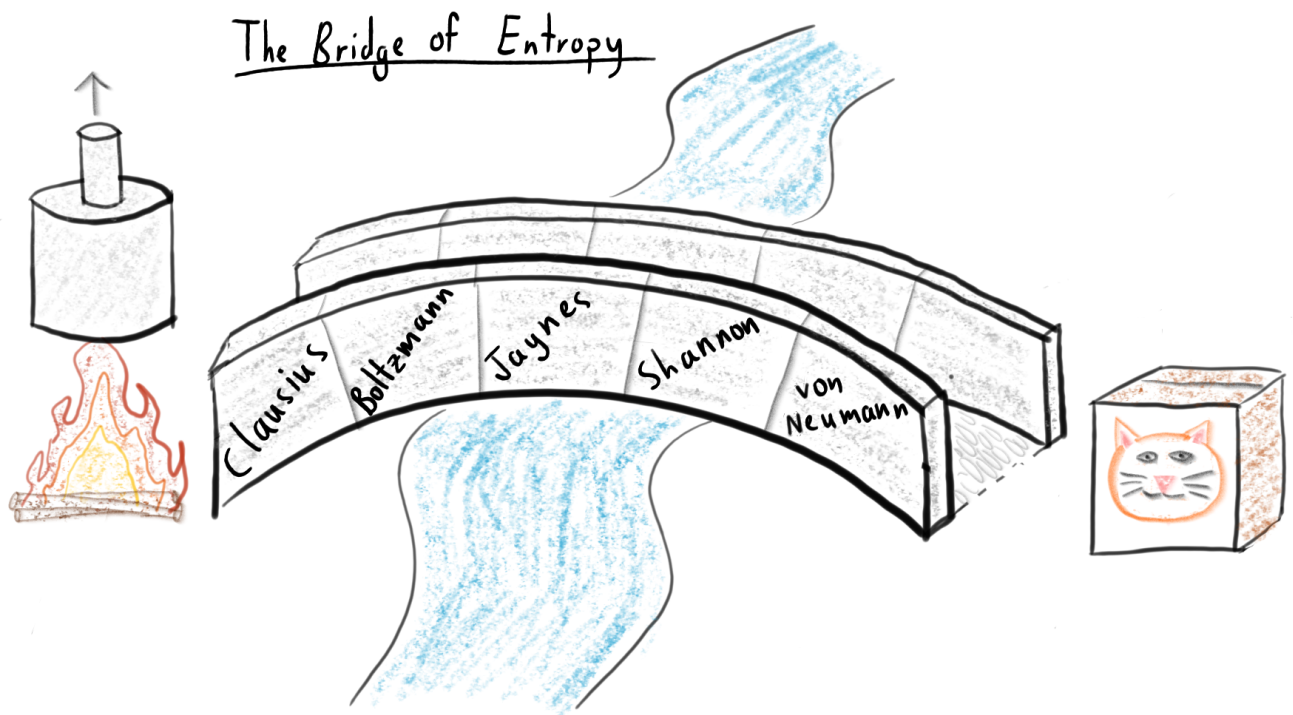
Clausius → Boltzmann

The first explicit appearance of entropy happened to Clausius, its namer, in 1865. Starting with his famous (in)equality for a closed reversible transformation at temperature T receiving a quantity of heat δQ,

$$\oint \frac{\delta Q}{T} \geq 0. \tag{1.1}$$

As the above holds for any cycle, a line integral over part of this cycle stops being path dependant (as one could just imagine a cycle that includes this line integral) allowing us to identify a state function satisfying,

$$dS = \frac{\delta Q}{T}. \tag{1.2}$$

This is our first encounter with the 'Second law' of thermodynamics where we have used d/δ to indicate a path independent/dependent variable change. Namely, this equality between entropy S and heat Q holds only for the 'perfect' transfer of said heat into the system. However, Clausius's statement tells us that such a perfect transfer of heat into the system is fictitious and in reality there will be *"...some other change, connected within..."* the system [59] that results in this equality becoming an inequality,

$$dS > \frac{\delta Q}{T}. \tag{1.3}$$

However, the entropy in the above equations is appearing merely as a mathematical construct. The question *"What is it?"* is still left unanswered. Thankfully Boltzmann was inspired to calculate the phase volume W of an ideal classical gas (one that obeys the ideal gas law) composed of N particles in volume V, for which the energy lies in $(E, E + dE)$,

$$W = \int_R d^3 x_1 \cdots d^3 x_N d^3 p_1 \cdots d^3 p_N = CV^N E^{\frac{3N}{2}-1} dE, \tag{1.4}$$

where R is the region of integration for coordinates within the volume, C is a constant independent of V, E and the momenta lie within the energy range $(E, E + dE)$. Starting from the entropy function identified by Clasius in (1.2) one can form the entropy of an ideal gas [150, Ch.4],

$$S(V, T) = \frac{N}{N_A} C_\mu \log T + N k_B \log V + constant \tag{1.5}$$

where, N_A is Avogadro's number and the *constant* is independent of T and V. Assuming this ideal gas is monatomic we can substitute in the following identities $C_\mu = (3/2)N_A k_B$ and $E = (3/2)Nk_B T$. This results in an entropy which has the same energy and volume dependence as the phase volume calculated by Boltzmann above, therefore up to an additive constant independent of T and V we arrive at his famous equation,

$$S = k_B \log W. \tag{1.6}$$

This remarkable result takes the mathematical construct of entropy derived from phenomenological macroscopic thermodynamics and explicitly connects it to the underlying phase volume W, occupied by all of the microstates compatible with the emergent macroscopic quantities.

Boltzmann $\rightarrow$ Jaynes

Let us start by thinking more deeply about the above entropy and its relationship to the underlying microstates of the system. First of all we can use equation (1.6) to receive some discernment into the aforementioned Second Law of thermodynamics, why does the entropy of a system tend to increase as indicated by the dynamics discussed by Clausius?

The Boltzmann equation tells us that the entropy of any macroscopic state is a measure of the phase volume occupied by all microstates compatible with that macrostate. Say a closed system has undergone some dynamics and is left with a choice between two distinct macrostates A and B, each of which has some set of compatible microstates. If, lets say $S_A < S_B$ why would the system have a greater preference for B rather than A? Is it because the microstates within B are somehow more preferred by the state than those in A? Not at all, there are just more of them to choose from, hence a system will tend to choose the macrostate with the higher entropy[1]. A simple example of the relationship between micro and macro states is illustrated in Figure 1.1.

Let us now go full circle and try to derive some thermodynamics from our newfound understanding of micro/macrostates. Say we have some system of identical particles, the only restriction on said system is that the total energy E, remains constant. We can therefore express said energy as,

$$E = \sum_{i=0}^{k} n_i \epsilon_i, \tag{1.7}$$

---

[1]Note that a system may well accidentally choose one of the less numerous microstates in the macrostate A of the system, therefore resulting in $dS < 0$, this apparent "breaking" of the law is studied within the framework of stochastic thermodynamics and will feature in section 1.4.
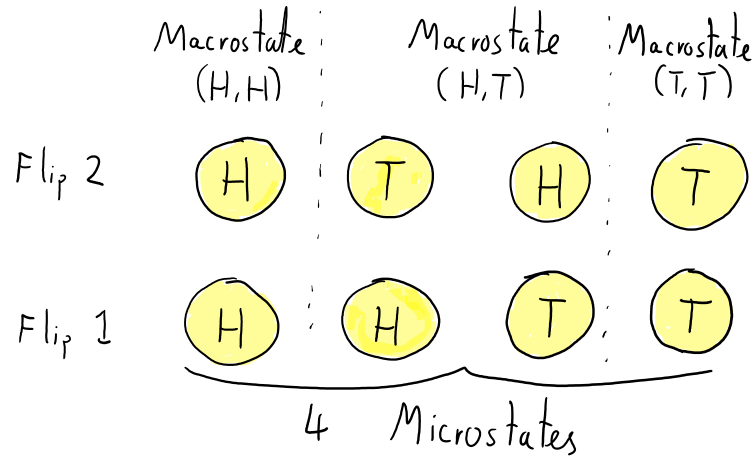
Figure 1.1: The relationship between the micro and macrostates of flipping a coin twice. Note that there are twice as many ways (microstate's) of forming macrostate (H,T) compared to the other two. Meaning this macrostate's phase volume, hence entropy, will be larger and therefore preferred in the *choice* dynamics of the system.

where n_i is the number of particles having energy ϵ_i. With the total number of particles N being

$$N = \sum_{i=0}^{k} n_i. \tag{1.8}$$

Let us now return to the phase volume, or number of microstates W, in Boltzmann's famous equation (1.6). How do we go about identifying said microstates of a system of identical particles? A microstate of a system refers to a description of the system in which the state of every individual particle is specified, like the Heads/Tails of the coins in Figure 1.1. Assuming that every one of the N particles were perfectly distinguishable there would be $N!$ unique ways to distribute them. However, these are identical particles and the only degree of freedom we have to differentiate them is their energy ϵ_i, within which there are $n_i!$ completely identical distributions per i, these distributions have to be divided out of the total number of permutations $N!$ Taking all of this into account we calculate the total number of microstates to be,

$$W = \frac{N!}{n_0! n_1! \cdots n_k!}. \tag{1.9}$$

The entropy of the state we have described is therefore proportional to the number of microstates above (1.6). I am labouring this point for the following reason; the entropy derived above is the one an observer would apply to a system if said observer had no information which suggested that the identical particles preferred one energy type ϵ_i, over another. Therefore said observer could do no better than to assume that any of the above

microstates, compatible with the emergent macroscopic quantities that they can see and measure (energy E, total particle number N), are all equally likely. This is known as the *equal a priori probabilities* postulate and it allows us to write down the probability of being in any one microstate as $p = 1/W$.

What does this assumption of maximum ignorance tell us, or indeed not tell us, about a system in equilibrium with its environment? Say we again have a System S with some set of energy levels $\{E_i\}$, in contact with a Bath B, the total combined energy of system and bath being E_{total}. Using the equal a priori probabilities postulate our best guess for the probability of the system being in microstate i is,

$$p_i = \frac{W_B(E_{\text{total}} - E_i)}{\sum_i W_B(E_{\text{total}} - E_i)}, \qquad (1.10)$$

where $W_B(E_{\text{total}} - E_i)$ is the number of microstates in the bath compatible with the system being in microstate i, and the denominator being the total number of microstates compatible with any system energy. Using Boltzmann's entropy formula equation (1.6), Taylor expanding the resultant logarithms and using the statistical mechanics formula $\partial S_B / \partial E_{\text{total}} = 1/T$, we arrive at the famous Boltzmann distribution,

$$p_i = \frac{1}{Z} e^{-\frac{E_i}{k_B T}}, \qquad (1.11)$$

where $Z = \sum_i e^{-\frac{E_i}{k_B T}}$ is the partition function[2].

The Boltzmann distribution is therefore an observer's *best guess* for the distribution of the system in equilibrium with its environment. Namely when they can observe the macroscopic quantities such as total energy E_{total}, but can only guess at the underlying microstate configuration. To quote Jaynes from his seminal paper [133]:

"There is nothing in the general laws of motion that can provide us with any additional information about the state of a system beyond what we have obtained from measurement."

---

[2]Note that the summation here is over the microstates of the system as opposed to the energies of the system, this can be converted by including the degeneracies $\{n_i\}$, see [162, Sec. 2.5]. This discrepancy between summing over microstates/energy levels will be discussed below.

Boltzmann $\rightarrow$ Jaynes $\rightarrow$ Shannon

We start with probably the most celebrated equation in information theory, the Shannon entropy,

$$H\left(\{p_i\}\right) = -K\sum_i p_i \ln p_i, \tag{1.12}$$

where $\{p_i\}$ is a set of probabilities and K is some constant. This quantity's foundational importance becomes clear in the famous *source coding* theorem of Shannon [209], which states that a source X emitting n random outputs can be compressed to length $nH\left(X\right) + \mathcal{O}(n)$, and restored to the original emission with high probability.

But what does this quantity have to do with the Boltzmann entropy from thermodynamics? We can investigate this question by looking at the maximisation of the Shannon entropy, such a maximisation can be solved using the well known method of Lagrangian multipliers. Our restrictions for this maximisation of a vector state $\mathbf{p} = (p_1, p_2, \dots, p_d)$ of dimension d being its normalisation $\sum_i p_i = 1$ and average energy $\sum_i p_i E_i = \bar{E}$, leading to the system of equations,

$$\frac{\partial}{\partial p_i}\left[-K\sum_i p_i \ln p_i - \lambda_1 \sum_i p_i E_i - \lambda_2 \sum_i p_i\right]_{p_i = \tilde{p}_i} = 0, \tag{1.13}$$

The probabilities which solve this being,

$$\tilde{p}_i = \frac{1}{Z}e^{-\lambda_1 E_i/K} \qquad \text{where} \qquad Z = \sum_i e^{-\lambda_1 E_i/K}, \tag{1.14}$$

where the Lagrangian multiplier λ_1 and constant K can be identified with the inverse temperature β showing that the state which maximises this entropic function is the Gibbs state $\boldsymbol{\gamma} = \frac{1}{Z}\left(e^{-\beta E_1}, e^{-\beta E_2}, \dots, e^{-\beta E_d}\right)$ whose probabilities exactly follow the Boltzmann distribution[3].

This is indeed the question asked by Jaynes to argue for his principle of Maximum Entropy (MaxEnt) [133]. Jaynes showed that the thermodynamic entropy of Boltzmann (1.6) emerges identical to the information-theoretic (Shannon) entropy, therefore creating a foundational connection between the two topics of thermodynamics and information theory. In particular how the maximisation of the Shannon entropy leads to the thermodynamic distribution calculated by Boltzmann through the equal a priori probabilities postulate, Jaynes's epistemological justification of this maximum entropy principle being:

---

[3]A similar derivation with more constraints leads to the derivation of the grand canonical distribution.

"It is uniquely determined as the one which is maximally non-committal with regard to missing information".

This implicit link between thermodynamics and the identification of the knowledge of some observer is something that will be frequently revisited in this thesis.

Shannon → von Neumann

Before we discuss how von Neumann used thermodynamics to motivate the construction of quantum theory and his entropic namesake, we must first define said construction. Introduced almost simultaneously by Dirac in 1930 [74] and von Neumann in 1932 [177], this can be summarized as the following:

- A physical system is described in a complex Hilbert space $\mathcal{H}$.

- The state of the quantum system can either be a pure state $|\psi\rangle$ (or ray) in this Hilbert space, or some density matrix describing a mixture of pure states $\sigma = \sum_i p_i |\psi_i\rangle\langle\psi_i|$.

- The observables of a quantum system are self-adjoint operators A on $\mathcal{H}$, the expectation value of which being $\mathrm{Tr}\,[\sigma A]$.

One of the things that von Neumann attempted to motivate in his treatise was an entropic quantity associated to this statistical operator σ. The way in which he chose to do this was via a thermodynamic thought experiment, whose stages I have detailed in Figure 1.2.

Let us consider an ideal gas of N particles in a box B. We first assume that said gas can be described by one of these aforementioned density matrices $\rho = \lambda |\psi_1\rangle\langle\psi_1| + (1-\lambda)|\psi_2\rangle\langle\psi_2|$ which is a mixture of two different states $|\psi_1\rangle, |\psi_2\rangle$ which exist inside this box, of which there are λN particles in state ψ_1 and $(1-\lambda)N$ particles in state ψ_2. We then make an important assumption that if ψ_1, ψ_2 are orthogonal to each other, then there exists some semi-permeable wall which could address them individually, letting one particle type pass though whilst prohibiting the movement of the other. We then let some experimenter perform the following protocol:

1. The box B containing the aforementioned system in the state ρ has a volume V and is at temperature T. The entropy of the initial gas, via the *extensivity* property, can be written as $S(\psi_1, \lambda N) + S(\psi_2, (1-\lambda)N)$, where the entropy is calculated at said volume and temperature.

2. An additional box of equal volume, called B' is attached to the left hand side of box B with a impregnable wall separating them.

3. The experimenter then inserts, just to the right of this impregnable wall, an additional semi permeable wall which is permeable to ψ_1 particles only. And on the right hand side of box B another wall which only lets through ψ_2 particles.

4. The experimenter then slowly moves both the impregnable wall and the rightmost wall left at the same speed, hence doing no work against the gas pressure, but separating the ψ_1 particles into Box B' whilst leaving the ψ_2 particles in Box B.

5. The semi permeable walls are then replaced with impregnable ones by the experimenter. The particles have now been separated without any work being done, change in temperature or heat.

6. The experimenter then isothermally compresses Box B' containing ψ_1 particles to volume λV and Box B containing ψ_2 particles to volume $(1 - \lambda)V$. The work required for the Boxes compression (B' in this example) being calculated using the ideal gas law, $\int_V^{\lambda V} P dV = \lambda N k_B T \ln(\lambda)$ where as the temperature and thus the internal energy remain constant, we extract an equivalent amount of heat and via equation (1.2) entropy.

7. Finally, the experimenter mixes the ψ_1, ψ_2 gasses to obtain a new gas mixture σ of N particles at volume V, temperature T, whose entropy we denote as $S(\sigma, N)$.
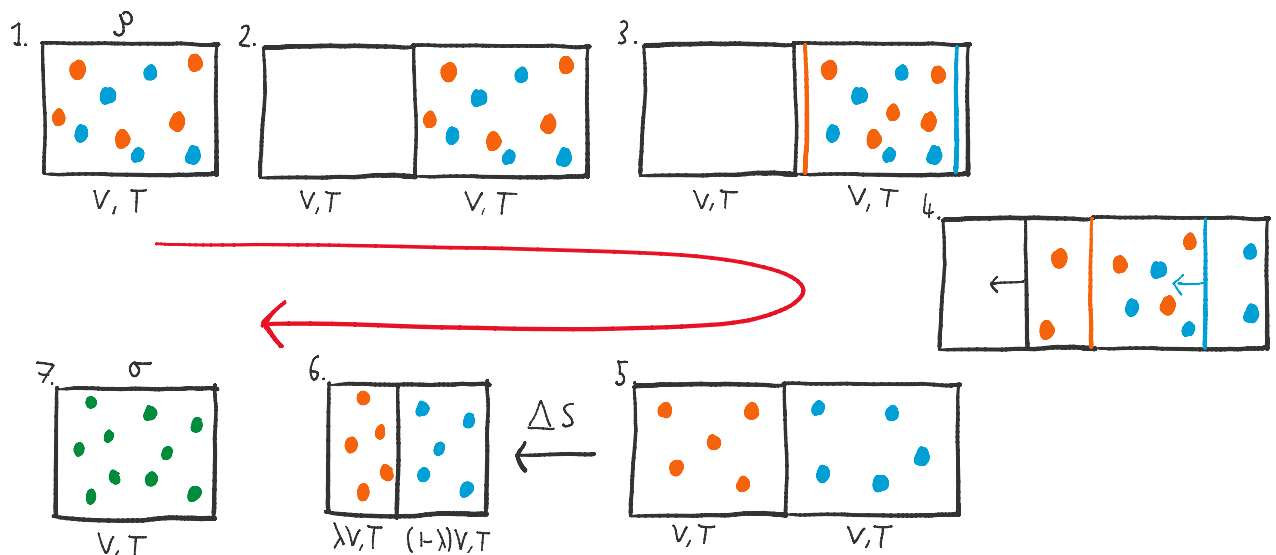
Figure 1.2: Diagram showing Von Neumann's original thought experiment. The seven separate stages being detailed above.

Now the only change in entropy from the initial state ρ to the final state σ was via the isothermal compression, allowing us to write the following equality,

$$S(\sigma, N) = S(\psi_1, \lambda N) + S(\psi_2, (1 - \lambda)N) - k_B N \left(\lambda \log \lambda + (1 - \lambda) \log(1 - \lambda)\right). \quad (1.15)$$

We know that the entropies of the ideal gasses are proportional to N (see equation (1.5)) therefore,

$$S(\sigma) = \lambda S(\psi_1) + (1 - \lambda)S(\psi_2) - k_B \left(\lambda \log \lambda + (1 - \lambda) \log(1 - \lambda)\right). \quad (1.16)$$

We can then notice that instead of there being two different particles in the mixture ψ_1, ψ_2 there is nothing stopping the experimenter performing the above protocol over some many component mixture, as long as all of those components are orthogonal,

$$S(\sigma) = \sum_i \lambda_i S(\psi_i) - k_B \sum_i \lambda_i \log \lambda_i. \quad (1.17)$$

As the entropy of a pure state is zero[4] we arrive at the formula,

$$S(\sigma) = -k_B \sum_i \lambda_i \log \lambda_i., \quad (1.18)$$

or in its more recognisable form,

$$S(\sigma) = -k_B \operatorname{Tr}[\sigma \log \sigma], \quad (1.19)$$

showing von Neumann's entropy to be consistent with thermodynamic entropy. Completing this entropic bridge 1.2 between thermodynamics and quantum mechanics. The path followed here between the two fields is by no means the only one connecting them but seemed to be the one best placed to motivate the rest of this thesis.

1.3 Particles and thermodynamics

One of the most important revolutions in thermodynamics, possibly the most important one, was its merging with statistical mechanics. This gave thermodynamics the power to describe how a system's macroscopic properties (temperature, pressure etc..) emerged from its underlying microscopic configuration, these microscopic configurations being the states and interactions of the particles comprising our thermodynamic state. Therefore

---

[4]A different proof is given by von Neumann is his original treatise [177] where he showed that all pure states can be reversibly transformed to one another, one can also follow the extension of Petz [190] who generalised the above for mixed states via a different route.

at a more fundamental level, thermodynamics could be argued to be a theory of particles rather than a theory of heat and work. This motivates an additional direction to approach quantum thermodynamics; what if instead of the particles comprising my state being well defined objects in phase space they are instead particle wavefunctions living in some complex Hilbert space?

In order to better understand the role of the particle in the preceding discussion on foundational thermodynamics, let us revisit the Gibbs' distribution (1.14) derived above to make an important point about moving between the microstate and particle picture.

The summation present in this distribution is a summation over the total energy E_i of each of the explicitly differentiable microstates i. What if we wanted to complete this summation not in terms of microstates but over the N identical particles comprising our classical gas? If we ignored interactions between particles, we can represent a N-particle microstate by the underlying particle configurations,

$$i = (j_1, j_2, \cdots, j_N)\,, \tag{1.20}$$

where j_1 is the *state* of particle 1, j_2 is the *state* of particle 2, etc. The total energy for that microstate i is therefore the sum of particle state energies,

$$E_i = \epsilon_{j_1} + \epsilon_{j_2} + \cdots + \epsilon_{j_N}, \tag{1.21}$$

where ϵ_{j_1} is the energy of particle 1 in state j_1 etc. Naively we can then rewrite the partition function in the N-particle Gibbs' distribution not as a summation over microstates but as a summation over the possible states of the particles,

$$
\begin{aligned}
Z = \sum_i e^{-\beta E_i} &= \sum_{j_1, j_2, \cdots, j_N} e^{-\beta\left(\epsilon_{j_1} + \epsilon_{j_2} + \cdots + \epsilon_{j_N}\right)} \\
&= \left(\sum_{j_1} e^{-\beta \epsilon_{j_1}}\right)\left(\sum_{j_2} e^{-\beta \epsilon_{j_2}}\right) \cdots \left(\sum_{j_N} e^{-\beta \epsilon_{j_N}}\right) \\
&= (Z_1)^N,
\end{aligned}
\tag{1.22}
$$

where $Z_1 = \sum_{j_1} e^{-\beta \epsilon_{j_1}}$. What are we implicitly assuming by writing the summations in this way? First of all, the assumption that each of these single state summations is the same can be motivated by the fact that all of the particles will have the same energetic spectrum, a reasonable assumption for identical particles. However in the step preceding this one, by splitting up the summation of states into products we are assuming that these summations are independent of one another, that each of these particles sits in its predesignated state without hopping into any of its neighbours. This is certainly true if

these state labels have some well defined meaning, like locations in a solid, but what if these particles states are not so well defined, like in a gas? We might accidentally count a configuration of particles states that resulted in two particles swapping places. In order to fix this overcounting classically we *'ad hoc'* divide out the extra configurations by $N!$, resulting in a partition function of the form,

$$Z = \frac{1}{N!}(Z_1)^N. \tag{1.23}$$

It is again worth thinking about what this means in Jaynes's epistemological language of thermodynamics. If an observer had the ability to follow the trajectories of said particles, that observer would have a well defined notion of each of the particles states in the gas, therefore would need no such $N!$ correcting factor in their partition function for the gas, whereas a standard observer of the 'truly' identical particle system would.

Is this always true? Is particle identity just a question of the resolving abilities of some observer?

The answer, for fundamental particles described by quantum mechanics, is in the negative. Some properties of quantum particles are independent of the resolving abilities of an experimenter. Let me motivate this statement and briefly introduce the effect responsible.

Elementary quantum particles[5] are known to belong to one of two classifications governing the behaviour of their wave function, being either bosonic (integer spin) or fermionic (half-integer spin) in nature 1.3. The restriction on their wavefunction being that for a system of identical (half) integer particles, upon swapping the position of any two particles, the wave function must remain (anti)symmetric. The foundations for this result lie in the spin-statistics theorem which emerges as a result of enforcing the physical invariance of the state under Lorentz transformations. Before we analyse what such a state of identical quantum particles would look like for one of the microstates discussed above let me first give some definitions regarding the representation theory of the symmetric group governing the behaviour of these (specialising to bosonic) wavefunctions.

Let $\mathcal{S}_N$ denote the symmetric group for N letters (particles), for some permutation $\pi \in \mathcal{S}_N$ we can define the representation of $\mathcal{S}_N$ on the full state space $\mathcal{H}^{\otimes N}$ where $\dim \mathcal{H} = d$ as,

$$P(\pi) := \sum_{j_1, j_2, \cdots, j_N} \left| j_{\pi^{-1}(1)}, \cdots, j_{\pi^{-1}(N)} \right\rangle \left\langle j_1, \cdots, j_N \right|. \tag{1.24}$$

---

[5]such classifications can also be made for some composite or quasi particles such as deuterium or phonons, but such classifications cannot be made in general.
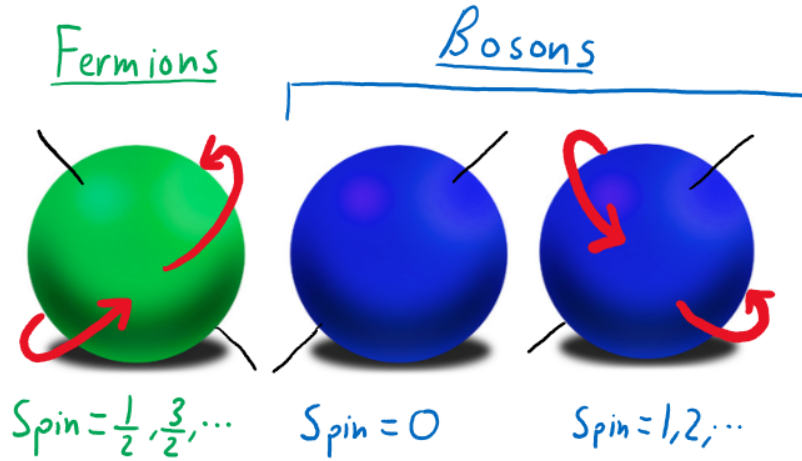
Figure 1.3: Elementary quantum particles are split into two distinct categories.

We can therefore properly define what we mean by the symmetric subspace $\mathcal{H}_+^{\otimes n}$,

$$\mathcal{H}_+^{\otimes N} := \left\{ |\psi\rangle \in \mathcal{H}^{\otimes N} : P(\pi) |\psi\rangle = |\psi\rangle \, \forall \, |\psi\rangle \in \mathcal{S}_N \right\}. \tag{1.25}$$

Using the $P(\pi)$ representation we can now define the orthogonal projector onto $\mathcal{H}_+^{\otimes N}$

$$P_+ := \frac{1}{N!} \sum_{\pi \in \mathcal{S}_N} P(\pi), \tag{1.26}$$

the necessary proofs for the above object being an orthogonal projector can be found here [113, p.3]. The above projector is the very object which ensures a wavefunction correctly describes a system of bosonic particles. Let us therefore look at microstate i described in (1.20), where each particle has its own state j,

$$|\psi\rangle_i = |j_1, j_2, \cdots, j_N\rangle. \tag{1.27}$$

However, now we are defining a system of bosonic particles *we must* project onto the symmetric subspace,

$$\begin{aligned} P_+ |\psi\rangle_i &= P_+ |j_1, j_2, \cdots, j_N\rangle \\ &= \frac{1}{N!} \sum_{\pi \in \mathcal{S}_N} \left| j_{\pi^{-1}(1)}, \cdots, j_{\pi^{-1}(N)} \right\rangle, \end{aligned} \tag{1.28}$$

like (1.23) we have arrived at an identical particle state where the configurations have been divided out by $N!$ However, remember in the classical example we included such a correcting factor '*ad hoc*' when such state labels describing the particles became ill defined for some observer and we needed to avoid overcounting. This is explicitly dependant on

the physical scenario and the resolving abilities of the experimenter in question. However, for the system of bosonic particles, this $N!$ is purely a result of the particles nature.

In chapter 4 we explore the result of this symmetrisation in the context of the historical thermodynamic paradox concerning indistinguishability, "The Gibbs Paradox" [97]. Finding that such an observer independent property of the particle wavefunction can provide you with a thermodynamic advantage in the quantum case.

1.3.1 Particle entanglement

In this section I want to introduce a property of these aforementioned (anti)symmetrised states that has generated a huge amount of discussion and controversy over the years. But before I do I first need to introduce a more foundational feature of quantum mechanics, 'entanglement'.

Say we have the following Hilbert space structure within which our states live,

$$\mathcal{H}_A \otimes \mathcal{H}_B, \tag{1.29}$$

the subscripts A, B may have some operational interpretation as distant labs or just be defining different subsystems. A general (pure) state $|\Psi\rangle_{AB}$ which lives in this joint space can be written as,

$$|\Psi\rangle_{AB} = \sum_{i,j} a_{ij} |\psi_i\rangle_A \otimes |\phi_j\rangle_B, \tag{1.30}$$

where $|\psi_i\rangle_A$ forms a basis for $\mathcal{H}_A$ and $|\phi_j\rangle_B$ for $\mathcal{H}_B$. We then say a state is entangled if for any i, j the vector $a_{ij} \neq a_i^A a_j^B$ prohibiting the state from being written as a product,

$$|\Psi\rangle_{AB} = \sum_{i,j} a_i^A |\psi_i\rangle_A \otimes a_j^B |\phi_j\rangle_B$$
$$= |\tilde{\psi}\rangle_A \otimes |\tilde{\phi}\rangle_B. \tag{1.31}$$

What does this actually mean for the state in question $|\Psi\rangle_{AB}$? Foundationally, it means that even if you had the best possible knowledge of the parts of the state which lived in $\mathcal{H}_{A,B}$ separately, you would not have the best possible knowledge of the whole system. There is some extra information 'entangled' between the two systems. In addition to being conceptually interesting, we can imbue the state with some operational interpretation. Say that the Hilbert spaces A, B corresponds to two distant separated labs, the PI's of these labs we often call Alice and Bob, who can only communicate via some classical

communication channel (can only exchange classical states). It turns out that sharing an entangled state allows Alice and Bob to perform information based tasks that would be impossible if the state they shared were not entangled. To name a few; quantum teleportation [31], superdense coding [34] and quantum cryptography [79], all of which have now been experimentally demonstrated. A modern approach to entanglement theory often takes place within the context of a resource theory, introduced in section 1.5, where the operational description above, known as Local Operations and Classical Communication (LOCC), helps us to both mathematically quantify and measure entanglement. In chapter 2 we utilize this construction to quantitatively investigate how shared entanglement between two parties may provide an advantage in the process of local work extraction.

Now we have an understanding of what it means for a state to be entangled let us return to the bosonic particle states introduced above. Say that in our spatially separated labs, both Alice and Bob are in possession of a bosonic particle, in this case with no internal degree of freedom. Naively we would write such a state (in first quantisation) as,

$$|\Psi\rangle_{AB} = |A\rangle_1 |B\rangle_2 \,, \tag{1.32}$$

where Alice is in possession of particle 1 and Bob in possession of particle 2. However, as these particles are bosonic, they must live in a a symmetric Hilbert space $\mathcal{H}_+$, therefore must remain invariant under the action of the permutation operator $P(\pi)$ from equation 1.24. After symmetrising the above state it appears as,

$$|\Psi\rangle_{AB} = \frac{|A\rangle_1 |B\rangle_2 + |A\rangle_2 |B\rangle_1}{\sqrt{2}}. \tag{1.33}$$

Alice and Bob seem, according to our above requirements, to have acquired an entangled state between particle 1 and 2, the above state cannot be written as a tensor product between 1 and 2. What does this mean, can Alice and Bob now perform some superdense coding task or teleport states to one another as they could do with a normal entangled state? The answer to this question is unfortunately in the negative and can be seen by looking more closely at what these particle labels $1, 2$ actually correspond to. Unlike a normal entangled state where the subscript refers to either local labs or some well defined subsystem, particle labels $1, 2$ are *by definition completely indistinguishable* therefore individually unaddressable, making any normal entanglement based task impossible.

Is this the end of story for particle entanglement, should we completely dismiss the apparent entanglement in the above state? Although it is fair to rule out the above state as being conventionally entangled we should not go as far as to dismiss the property it displays as unphysical. For example, such states are known to be incredibly important in quantum metrology [191] as they represent spin-squeezed states, and the emergent

properties of Bose-Einstein condensates are very much physical manifestations of symmetrisation. Where does this leave us with quantifying the above property of identical particle states?

In chapter 5, we provide such particle entanglement with a consistent theoretical description as a quantum resource and demonstrate it is precisely the resource required for the activation of an identical particle state into usable entanglement. We also apply our tools to an experimental implementation with Bose-Einstein condensates which leads to the first ever quantitative estimation of identical particle entanglement.

1.4 Fluctuation relations

The framework of stochastic thermodynamics [207] lends itself to studying the thermodynamics of small systems. This approach involves treating the variables of thermodynamics, such as heat or work, as stochastic variables that can be fully characterised by their probability distribution. This approach has lead to the development of the fluctuation relations that describe the out of equilibrium response of a thermodynamic system.

In particular interest of this approach is studying the probability distribution of extractable work. Consider a quantum system with a time-dependant Hamiltonian $H(\lambda(t))$ where $\lambda(t)$ is the work parameter [6]. The system is prepared by allowing it to equilibrate with a heat bath at temperature β where the work parameter is fixed at λ_i. The initial state of the system is therefore,

$$\gamma_\beta(\lambda) = \frac{e^{-\beta H(\lambda_i)}}{\text{Tr}(e^{-\beta H(\lambda_i)})} = \frac{e^{-\beta H(\lambda_i)}}{Z_\beta(\lambda_i)}. \tag{1.34}$$

The system then undergoes the following protocol:

1. At $t = t_i$, while the system is in thermal equilibrium, it is projected onto the energy eigenbasis of the initial Hamiltonian

$$H(\lambda_i) = \sum_n E_n(\lambda_i)|\psi_n\rangle\langle\psi_n|, \tag{1.35}$$

where $|\psi_n\rangle$ is the initial eigenstate of the initial Hamiltonian with eigenvalue $E_n(\lambda_i)$.

---

[6] At this stage I do not want to become sidetracked with any particular definition for extractable work, this will be introduced in more detail in section 1.5.1.

2. The system then evolves under the unitary operator,

$$U(t_f, t_i) = T_{\rightarrow} \exp\left(-i \int_{t_i}^{t_f} dt' H(\lambda(t'))\right), \tag{1.36}$$

where $T_{\rightarrow}$ is the time ordering operator. This process is allowed to bring the system out of equilibrium.

3. Finally at $t = t_f$ the system is then projected onto the energy eigenbasis of the final Hamiltonian

$$H(\lambda_f) = \sum_m E_m(\lambda_f)|\phi_m\rangle\langle\phi_m|, \tag{1.37}$$

defined in a analogous way to $H(\lambda_i)$.

Following the above protocol, it is easy to define the joint probability of obtaining $E_n(\lambda_i)$ for the initial measurement followed by $E_m(\lambda_f)$ for the second,

$$
\begin{aligned}
p(n, m) &= p_n |\langle\phi_m|U(t_f, t_i)|\psi_n\rangle|^2 \\
&= \frac{e^{-\beta H(\lambda_i)}}{Z_\beta(\lambda_i)} |\langle\phi_m|U(t_f, t_i)|\psi_n\rangle|^2.
\end{aligned} \tag{1.38}
$$

As the evolution is unitary, there is no increase in entropy of the system and hence work is defined as a discrete value that is equal to the difference in initial and final energies $W = E_m(\lambda_f) - E_n(\lambda_i)$. Then by weighting these values of work according to the probability in equation (1.38) we can form the work distribution,

$$P_F(W) = \sum_{n,m} p(n, m)\delta(W - [E_m(\lambda_f) - E_n(\lambda_i)]), \tag{1.39}$$

where δ is the Dirac delta function and F denotes the forward protocol. In order to study the fluctuation of the system, we now introduce a backward process. This is simply the reverse of the steps defined above with the evolution being time-reversed, $\Theta U(t_f, t_i)\Theta$, where Θ is the anti unitary time reversal operator. By taking the ratio of these forward and backward protocols, it is easy to form the Tasaki-Crooks relation [64],

$$\frac{P_F(W)}{P_B(-W)} = e^{\beta(W - \Delta F)}, \tag{1.40}$$

where ΔF is given by the well known expression for change in free energy from statistical

mechanics,

$$\Delta F = \frac{1}{\beta} \ln\left(\frac{Z_\beta(\lambda_i)}{Z_\beta(\lambda_f)}\right). \tag{1.41}$$

From the Tasaki-Crooks relation (1.40), by rearranging and integrating over work we get the Jarzynski equality,

$$\int dW P_F(W) e^{-\beta W} = \langle e^{-\beta W} \rangle = e^{-\beta \Delta F}. \tag{1.42}$$

Employing Jensen's equality on Jarzynski's equation [132] allows one to form an expression of the second law of thermodynamics $\langle W \rangle \geq \Delta F$. By taking the logarithm of both sides of the Tasaki-Crook's relation (1.40) and integrating over the forward probability distribution one can find an expression for this dissipated work,

$$\langle W \rangle_{\text{diss}} = \beta(\langle W \rangle - \Delta F) = K(P_F(W)||P_B(-W)), \tag{1.43}$$

where K is the Kullback Leiber divergence. This corresponds physically to an internal entropy change that would be detected as a heat source if placed in an ideal thermal bath at the end of the protocol. It has also been shown [69], that this dissipated work can be expressed as the quantum relative entropy,

$$\langle W \rangle_{\text{diss}} = S(\rho||\gamma_\beta(\lambda_f)), \tag{1.44}$$

where $\rho = U(t_f, t_i)\gamma_\beta(\lambda_i)U^\dagger(t_f, t_i)$ is the out of equilibrium state at the end of the protocol.

The above results have also been investigated in the lab with a number of experiments claiming to see the emergence of fluctuation theorems in microscopic systems [60, 164, 246, 11]. With some of these experiments utilizing systems such as RNA molecules, more recent discussion and debate have looked into the application of these microscopic fluctuation theorems in biological systems [227], an intriguing crossover between fields.

In chapter 3 using the underlying structure these fluctuation relations have in majorisation theory [8], we study the possible role these fluctuation relations may have in quantum phenomena, in particular quantum coherence, introduced in section 1.5.2.

1.5 Resource theories

A resource is defined as anything that possesses a property which in some way can be understood to be useful. A resource theory is a mathematical framework which allows

you to ask fundamental questions about the creation, activation and processing of said resource. A generic resource theory is composed of and defined by two key elements:

- The *free states*, those that can be created at no resource cost [7] and used for 'free'

- The *free operations* are the operations that can be performed at no cost [8]

The generality of such a construction has had a plethora of applications in quantum mechanics where the identification and understanding of quantum resources [57] (quantum phenomena which can be understood as useful) is one of the central motivations for the field. Given the above definitions one can ask: If an experimenter has access to a quantum state ρ, what states can they reach/how much resource can they extract while having access to as many free states and being able to perform as many free operations as they wish?

Such a question almost perfectly reflects the sorts of questions we are looking to answer in thermodynamics. For example, given a state in contact with a heat bath, how much useful energy (work) can be extracted globally while preserving energy. This has lead scientists to attempt to reformulate thermodynamics using the construction of a resource theory.

1.5.1   Resource theories of thermodynamics

Definition 1. *The resource theory of Thermal Operations (TO) is defined by the following two components [41]*

- Free states*: the set of all Gibbs states at a fixed temperature $\beta = 1/k_B T$, i.e $\gamma_R(\beta) = e^{-\beta H_R}/Z_R$ where $Z_R = \mathrm{Tr}\left[e^{-\beta H_R}\right]$ and H_R is an arbitary Hamiltonian;*

- Free operations*: partial traces and energy-preserving unitaries.*

Based on this definition, one can define the state transformation under TO using a Stinespring dilation. Namely, given a state ρ_S with Hamiltonian H_S, you can define the

---

[7] One is tempted to define this as states that possesses no resource, however there does exist states in some resource theories which cost resources to make but from which no resource can be extracted [126].

[8] The defining of free operations for each resource is often open to much debate/discussion and varies greatly depending on the degree to which the construction is required to be operational.

set of TO as,

$$\mathcal{E}_{\mathrm{TO}}\left(\rho_S\right) := \mathrm{Tr}_R\left[U_{SR}\left(\rho_S \otimes \gamma_R^{\otimes n}\right)U_{SR}^{\dagger}\right], \qquad \left[U_{SR}, H_S \otimes \mathbb{1}_R^{\mathrm{total}} + \mathbb{1}_S \otimes H_R^{\mathrm{total}}\right] = 0,$$

$$(1.45)$$

where H_S and $H_R^{\mathrm{total}} = \sum_i^n H_R^i$ are non-interacting Hamiltonians of the system and reservoir. Note also that $\mathcal{E}_{\mathrm{TO}}(\gamma_S) = \gamma_S$ for all U ensuring we cannot create infinite energy for free.

So far we have seen the quantity 'thermodynamic work' appear as a consequence of the first law of thermodynamics, relating the change in internal energy of a system to the heat added. However, an operational understanding of this definition is something which is still widely discussed and researched by the community, especially when one wants to talk about microscopic systems [89].

One way to understand this process and make it more explicit is to consider extracting work as the charging process of a battery system while under the construction of TO,

$$\rho_S \otimes |0\rangle\langle 0|_W \overset{\mathrm{TO}}{\to} |1\rangle\langle 1|_W \,,$$

where $H_W = w\,|1\rangle\langle 1|_W$.

The amount of work you can extract from a system using this construction can be studied in different regimes. For example the many copy regime [41],

$$\lim_{n\to\infty}\left(\rho_S^{\otimes n} \otimes |0\rangle\langle 0|_W^{\otimes Rn} \overset{\mathrm{TO}}{\to} |1\rangle\langle 1|_W^{\otimes Rn}\right),$$

where $R = m/n$ is the rate n systems can get transformed into m systems. It can be shown from this definition A.1 the extractable work is,

$$W(\rho_S) := \frac{1}{\beta}D(\rho_S||\gamma_S)$$

$$= F(\rho_S) - F(\gamma_S), \qquad (1.46)$$

where $D(\rho||\sigma) = \mathrm{Tr}\left(\rho\log\rho - \rho\log\sigma\right)$ and $F(\rho) = \mathrm{Tr}[H\rho] - \frac{1}{\beta}S(\rho)$.

We immediately note that the relative entropy has previously appeared as dissipated work in the aforementioned section on fluctuation relations. However, it is important to note that it is presently not clear if the resource theoretic approach is comparable with the stochastic variable approach previously defined. The resource theoretic approach claims to fully quantify every step of the system by strongly defining what it classes as thermodynamically allowed operations, whereas the stochastic thermodynamic approach

has been criticised by not considering the full thermodynamic cost of its operations. It is hoped that with further investigation, for example this recent work [107], more links will emerge attempting to bridge the gap between these apparently disparate approaches to microscopic thermodynamics.

We can also study work extraction in the single-shot regime [127] where the transformation is exact,

$$\rho_S \otimes |0\rangle\langle 0|_W \xrightarrow{\text{TO}} \gamma_S \otimes |1\rangle\langle 1|_W.$$

Which gives the extractable work as,

$$
\begin{aligned}
W_{\min}(\rho_S) :=& \frac{1}{\beta} D_{\min}(\rho_S || \gamma_S) \\
=& F_{\min}(\rho_S) - F_{\min}(\gamma_S),
\end{aligned}
\tag{1.47}
$$

where $D_{\min}(\rho || \gamma) = -\ln\left(\text{Tr}\, \Pi_\rho \gamma\right)$ is the min relative entropy with Π_ρ being the projector onto the support of ρ in the energy eigenspace and $F_{\min}(\rho) := -\frac{1}{\beta} \ln \text{Tr}\, \Pi_\rho e^{-\beta H}$ is the single shot free energy.

You can also study work extraction in the single-shot regime [127] where the charging process can happen with some error,

$$\rho_S \otimes |0\rangle\langle 0|_W \xrightarrow{\text{TO}} \gamma_S \otimes \left((1-\epsilon)\,|1\rangle\langle 1|_W + \epsilon \rho'\right).$$

Which gives the extractable work as,

$$
\begin{aligned}
W_{\min}^\epsilon(\rho_S) :=& \frac{1}{\beta} D_{\min}^\epsilon(\rho_S || \gamma_S) \\
=& F_{\min}^\epsilon(\rho_S) - F_{\min}(\gamma_S),
\end{aligned}
\tag{1.48}
$$

where $D_{\min}^\epsilon(\rho || \gamma) = -\ln\left(\text{Tr}\, \Pi_\rho^\epsilon \gamma\right)$ is the min relative entropy with allowable error of failure ϵ and $F_{\min}^\epsilon(\rho) := -\frac{1}{\beta} \ln \text{Tr}\, \Pi_\rho^\epsilon e^{-\beta H}$. We define Π_ρ^ϵ by β–ordering the energy eigenspaces of ρ so that $p(E_1)e^{\beta E_1} \geq p(E_2)e^{\beta E_2}$ and summing from $p(E_1)$ up to some value l such that $\sum_{i=1}^{l} p(E_i) \leq \epsilon$ and $\sum_{i=1}^{l+1} p(E_i) \geq \epsilon$ giving Π_ρ^ϵ the following definition ,

$$
\Pi_\rho^\epsilon := \begin{cases} 0 & \text{for } i < l \\ \frac{\epsilon - \sum_{i=1}^{l} p(E_i)}{p(E_i)} & \text{for } i = l \\ 1 & \text{for } i > l. \end{cases}
$$

Note that in all three cases, the work extracted is independent of the battery system

used. Also, the unitary which performs this operation is dependant upon the system and bath state. All we know is that a unitary exists which performs the above work extraction protocols.

We can also make the following assumption to simplify TO. By making the system and bath Hamiltonian trivial $H_S = H_B = 0$, this simplifies TO such that any global unitary can be performed. The Gibbs state also changes into the maximally mixed state $\gamma \overset{H\to 0}{=} \frac{\mathbb{1}}{\dim(\mathcal{H})}$ where $dim(\mathcal{H})$ is the full dimension of the Hilbert space[9]. The extractable work also loses its β weighting,

$$
\begin{aligned}
D_{\min}(\rho||\gamma) &\overset{H\to 0}{=} -\ln\left(\operatorname{Tr}\Pi_\rho \frac{\mathbb{1}}{\dim(\mathcal{H})}\right) \\
&= -\ln\dim(\rho) + \ln\dim(\mathcal{H}).
\end{aligned}
\tag{1.49}
$$

We can therefore see that work can be extracted from any state that doesn't explore its full Hilbert space. We can therefore see that the Landauer bound of $\beta^{-1}\log 2$ can be recovered from the above assumption if you assume that your system is a pure state in a two level Hilbert space with a trivial Hamiltonian.

1.5.2   Resource theories of quantum coherence

Quantum coherence is an essential non-classical feature rooted in the foundations of quantum theory. By fixing a particular reference basis $\{|i\rangle\}_{i=1,\dots,d}$ of the d-dimensional Hilbert space $\mathcal{H}$ in which the quantum states of our system of interest live, coherence is simply visualised as the degree to which these states deviate from being diagonal in the chosen basis. Although elementary in its conception, quantum coherence incarnates the essence of superposition and is thus seen as the first step away from a fully classical description of a system, acting as a building block for more advanced phenomena such as entanglement in composite systems.

It is of no surprise therefore that quantum coherence plays a central role in a wide range of quantum technologies, such as metrology, sensing, communication, and imaging. The development of these quantum technologies has motivated the formalisation of quantum coherence as a physical resource within the mathematical framework of resource theories [216]. This has led, *inter alia*, to theoretical and experimental investigations of optimal protocols to distill or dilute quantum coherence, and more generally to manipulate and transform quantum states by means of suitably defined free operations unable to create coherence [3, 24, 242, 218, 173, 55, 54, 56, 254, 236, 245].

---

[9]Note that as per the defintion of work from [127] we restrict ourselves to finite dimensional spaces.

Let us introduce the basics of quantum coherence from a resource theoretic perspective, referring the reader to [216] for more details.

Definition 2. *The resource theory of quantum coherence is defined by the following two components [24]*

- Free states: *given a reference basis $\{|i\rangle\}_{i=1,...,d}$ of a d-dimensional Hilbert space $\mathcal{H}$, such as e.g. the computational basis, density matrices of the form $\sigma = \sum_i c_i |i\rangle\langle i|$, form the set $\mathcal{I}$ of incoherent (free) states [10].*

- Free operations: *Our minimum requirement for this set is $\Lambda(\mathcal{I}) \in \mathcal{I}$, the set of Maximally Incoherent Operations (MIO) [3].*

As with every quantum operation such a map can be represented by a dilated map of the system and an environment interacting via some unitary [180], much like the map described by TO 1.45. However, it was shown that the set of MIO do not admit a free dilation [54, 55, 165], namely that the joint unitary and/or environment state were themselves not incoherent.

One such possibility for restricting this class and fixing this issue is the set of Incoherent operations (IO) which are completely positive trace preserving maps Λ admitting an operator sum representation such that all Kraus operators $\{K_l\}$ map incoherent states into incoherent states, that is, $\Lambda(\rho) = \sum_l K_l \rho K_l^\dagger$, with $\sum_l K_l^\dagger K_l = \mathbb{1}$ and

$$\frac{K_l \sigma K_l^\dagger}{\mathrm{Tr}(K_l \sigma K_l^\dagger)} \in \mathcal{I}, \quad \forall \sigma \in \mathcal{I}. \tag{1.50}$$

This definition entails that IO cannot create coherence from an incoherent state, not even probabilistically.

Strictly incoherent operations (SIO) are a subclass of IO whose Kraus operators additionally satisfy [242]

$$\frac{K_l^\dagger \sigma K_l}{\mathrm{Tr}(K_l^\dagger \sigma K_l)} \in \mathcal{I}, \quad \forall \sigma \in \mathcal{I}. \tag{1.51}$$

This equivalently means that the results of measuring (in the reference basis) an output state after SIO do not depend on the coherence of the input state ρ [248],

$$\langle i|K_l \rho K_l^\dagger|i\rangle = \langle i|K_l \Delta(\rho) K_l^\dagger|i\rangle, \tag{1.52}$$

---

[10]For a composite system with Hilbert space $\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B$, the reference basis is taken as the tensor product of the reference bases of each individual subsystem, and the set $\mathcal{I}$ of incoherent states is defined accordingly.

where we have introduced the dephasing operation Δ, whose action is defined as

$$\Delta(\rho) := \sum_{i=1}^{d} |i\rangle\langle i|\rho|i\rangle\langle i|\,, \tag{1.53}$$

such an operation destroys, in a particular basis, the 'quantum' coherence present in a state. There are several monotones apt to quantify the degree of coherence of a quantum state ρ [216]. One example being the relative entropy of coherence C_{rel} [3, 24, 118, 231, 104], which takes the simple closed form

$$C_{\mathrm{rel}}(\rho) := S(\Delta(\rho)) - S(\rho)\,, \tag{1.54}$$

where $S(\rho) := -\mathrm{Tr}(\rho \ln \rho)$ is the conventional von Neumann entropy, that is used prominently in quantum information theory as well as in extensive thermodynamics. The relative entropy of coherence admits a valuable operational interpretation as it amounts to the distillable coherence under IO in an asymptotic setting [242] [11].

A maximally coherent state in a Hilbert space of dimension d can be written as a uniform superposition of the reference basis states,

$$\left|\phi_d^+\right\rangle := \frac{1}{\sqrt{d}}\sum_{i=1}^{d} |i\rangle\,, \tag{1.55}$$

and its coherence is given by $C_{\mathrm{rel}}(\phi_d^+) = \ln(d)$.

There are still plenty of open questions concerning the resource theory of coherence and the role it plays in quantum systems. For example, unlike the physically well defined LOCC structure for the free operations of entanglement there is no equivalent understanding for quantum coherence. Although works [248] have been able to motivate a class of operations (SIO defined above) with an operational interpretation, these have since been shown to be somewhat weak [149, 148], in terms of their ability to asymptotically distil pure bits of coherence from a state. A final answer on the '*correct*' set of free operations for quantum coherence is by no means an answered question.

Nevertheless with the progress made on the topic so far we can now investigate the role of quantum coherence in quantum technologies, many-body physics, biological transport, and relevant to this thesis, thermodynamics. In Chapter 3 we use the above resource

---

[11]The distillable coherence $C_{\mathrm{d}}(\rho)$ of a state ρ is defined as the maximum ratio R such that the conversion $\rho^{\otimes m} \to \left|\phi_2^+\right\rangle\left\langle\phi_2^+\right|^{\otimes mR}$ can be implemented by IO in the limit of many copies $m \to \infty$. Strictly speaking, the equality between distillable coherence and relative entropy of coherence holds when $\log_2$ is used instead of ln in the definition of entropy, as it is customary in information theory. In Chapter 3, we adopt instead natural logarithms to better emphasise the connection with thermodynamics, which means that in our notation we have $C_{\mathrm{d}}(\rho) = C_{\mathrm{rel}}(\rho)/\ln(2)$.

theory of quantum coherence to investigate the emergence of thermodynamic fluctuation like relations for pure quantum states under IO or SIO assisted by a coherence battery, that is, a storage device whose degree of coherence is allowed to fluctuate in the process.

Chapter 2

Assisted work distillation

2.1 Introduction

As discussed in the previous chapter, quantum thermodynamics represents a drive to understand the interplay of the two fundamental theories of thermodynamics and quantum mechanics. This may be approached from various disciplines such as open quantum systems [39, ch.1], stochastic thermodynamics [80] and information theory [103] all of which are utilising their respective tools to answer these fundamental questions. In this chapter, we will utilise the resource theoretic construction defined in section 1.5 to investigate thermodynamic transformations in a quantum information setting.

The task under investigation is that of *assisted work distillation*, see Fig. 2.1. Here, the process of work distillation is intended in a resource theoretic framework to be the asymptotic distillation of reference states with energy but no entropy by means of thermal operations, meaning that the *distillable* (or *extractable*) *work* can be quantified by how distinguishable a quantum state is from a Gibbs equilibrium state [41] — for other definitions of work in quantum thermodynamics see e.g. [103]. In the assisted scenario, two parties, Alice (A) and Bob (B), share many copies of a bipartite state ρ_{AB}. Between them their goal is to maximise the quantity of distillable work on Bob's subsystem. Alice may perform arbitrary quantum operations on her subsystem whereas Bob is restricted to thermal operations on his. By utilising correlations within ρ_{AB} and classical communication between the parties we demonstrate key features of Bob's distillable work.

In particular, we characterise the set of shared states which allow for local work distillation. We also demonstrate that for a protocol involving one-way communication between the parties explicit expressions for the local distillable work, which we dub the *work of*

assistance (in analogy with the *entanglement of assistance* [75]), can be derived both in the regularised and un-regularised scenarios. From these expressions we make use of two central results from quantum information theory to show that Alice performing global measurements over many copies of the shared state offers an explicit advantage over single copy measurements. We also show that this advantage disappears when the initial state is pure.

In addition to the *work of assistance* we also define the *work of collaboration*, defined as such to allow two-way communication between the parties and local Gibbs-preserving operations [82] on Bob's side. We show that by allowing this collaboration and the wider class of operations, the local distillable work can increase. We also demonstrate that for an initial pure state the *work of collaboration* may yield an increase in distillable work by an amount proportional to the entropy of Bob's subsystem $S(\rho_B)$, where $\rho_B = \text{Tr}_A [\rho_{AB}]$.
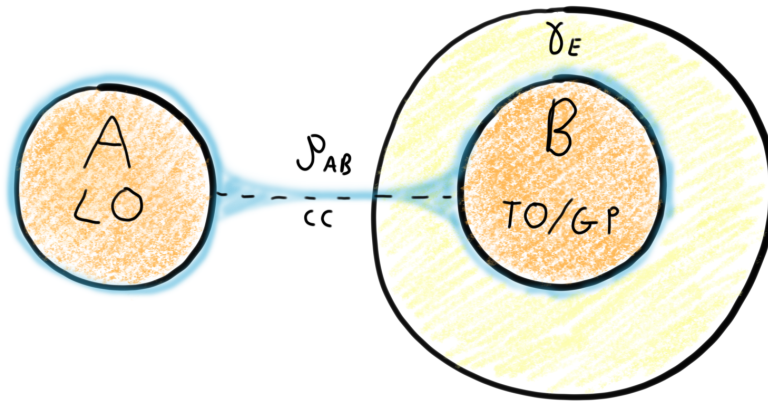
Figure 2.1: We investigate distillation of work from a quantum system B controlled by an observer, Bob, who is constrained to thermal operations or Gibbs-preserving operations, and is assisted by another party, Alice, who can perform arbitrary local operations on an ancillary system A and communicate classically with Bob. The *work of assistance* and the *work of collaboration* are defined and related to the correlations in the state ρ_{AB} shared by Alice and Bob.

It is important to consider the realm in which our results apply. Within the resource theoretical framework it is typical to consider resource inconvertibility in the asymptotic scenario. This is particularly pertinent for thermodynamics due to its equivalence to taking the thermodynamic limit, which suppresses the appearance of fluctuations.

2.2   Resource theories of thermodynamics

As introduced in section 1.5 the allowed TO for a quantum system S with Hilbert space $\mathcal{H}$ and Hamiltonian H_S are the completely-positive trace-preserving (CPTP) maps $\mathcal{E}$:

$\mathcal{L}(\mathcal{H}) \to \mathcal{L}(\mathcal{H})$ of the form

$$\mathcal{E}(\rho) = \mathrm{Tr}_E \left(U_{SE} \left(\rho_S \otimes \gamma_E \right) U_{SE}^\dagger \right), \tag{2.1}$$

where U_{SE} is an arbitrary unitary operation, acting jointly on the system S and a reservoir E, that commutes with the global Hamiltonian $[U_{SE}, H_S \otimes \mathbb{1}_E + \mathbb{1}_S \otimes H_E] = 0$, and $\gamma = Z^{-1} e^{-\beta H}$ denotes the Gibbs thermal equilibrium state at inverse temperature β and partition function Z. The joint unitary operations and partial trace define the *free operations* of the resource theory whereas the Gibbs states define the *free states*. By explicitly accounting for the resources used, the TO framework provides a general setting within which to study thermodynamic transformations, in particular the distillation of work.

In this setting, following [127] we define the distillable work from a system B in the state ρ_B as the maximum number RE such that the transformations $\rho_B^{\otimes n} \otimes |0\rangle\langle 0|_P^{\otimes [Rn]} \to |1\rangle\langle 1|_P^{\otimes [Rn]}$ are possible with TO at background inverse temperature β with asymptotically vanishing error. Here, referring to (2.1), we are considering a composite system S which consists of the principal system B with Hamiltonian H_B and a qubit battery P with Hamiltonian $H_P := E |1\rangle\langle 1|_P$, where E is a free parameter we are allowed to optimise over. In formula,

$$\begin{aligned} W(\rho_B) \;\; := \;\; \sup \Big\{ RE : \\ \lim_{n\to\infty} \inf_{\Lambda \in \mathrm{TO}} \left\| \Lambda \left(\rho_B^{\otimes n} \otimes |0\rangle\langle 0|_P^{\otimes [Rn]} \right) - |1\rangle\langle 1|_P^{\otimes [Rn]} \right\|_1 = 0 \Big\}. \end{aligned} \tag{2.2}$$

It follows from the main result of [41] (see appendix A.1 for an explicit derivation) that the distillable work defined in Eq. (2.2) equals the change in *free energy*:

$$W(\rho_B) \equiv \Delta F(\rho_B) = \frac{1}{\beta} S(\rho_B \| \gamma_B), \tag{2.3}$$

with $S(\rho \| \gamma) = \mathrm{Tr} \left(\rho \log \rho - \rho \log \gamma \right)$ being the relative entropy of athermality. Observe that $S(\rho \| \gamma)$ is monotonically non-increasing under TO.

A larger class of operations are Gibbs-Preserving (GP) operations; these are CPTP maps Λ that admit as their fixed point the Gibbs state at a given temperature, i.e. such that $\Lambda(\gamma_B) = \gamma_B$. The motivation behind this alternative framework that regards GP operations as free operations for thermodynamics, is that any non-GP operation, $\Lambda(\gamma) = \sigma \neq \gamma$, could be used to extract an arbitrarily large amount of work from $\sigma^{\otimes n}$ as $n \to \infty$. It can be clearly seen from (2.1) that TO are a subset of GP, and the inclusion is known to be strict [82].

2.3   Work of assistance

In this section we consider the case where Alice and Bob have access to the shared state ρ_{AB} and we allow one-way classical communication from Alice to Bob. This is similarly motivated as the recently studied 'conditioned thermal operations' [175]. Alice, whom operations are unrestricted, may perform on her subsystem the positive operator-valued measurement (POVM) $\{\Pi_{A,i}\}$, whose associated probabilities are $p_i = \mathrm{Tr}\,[\rho_A \Pi_{A,i}]$, whereas Bob is restricted to TO. Alice performing her measurement and communicating the outcome to Bob results in him having access to the ensemble $\{p_i, \tilde{\rho}_{B,i}\}$, where

$$\tilde{\rho}_{B,i} = \frac{1}{p_i}\,\mathrm{Tr}_A\left[(\Pi_{A,i} \otimes \mathbb{1}_B)\,\rho_{AB}\right]. \tag{2.4}$$

In the scenario we consider, Alice's goal is to help Bob to distil as much work as possible. From this train of thoughts we define our first quantity of interest, the *work of assistance*,

$$W_a^{B|A}(\rho_{AB}) := \max_{\{\Pi_{A,i}\}} \frac{1}{\beta} \sum_i p_i S(\tilde{\rho}_{B,i} \| \gamma_B), \tag{2.5}$$

where the maximisation is taken over the set of Alice's measurements (i.e. POVMs). Using convexity, we see that this quantity is lower bounded by $\frac{1}{\beta} S(\rho_B \| \gamma_B)$, which of course means that being assisted by Alice is generally no worse than having no assistance at all. Not only: as we show in appendix A.2, all states ρ_{AB} that exhibit some form of correlation, i.e. such that $\rho_{AB} \neq \rho_A \otimes \rho_B$ is not factorised, satisfy the strict inequality $W_a^{B|A}(\rho_{AB}) > \frac{1}{\beta} S(\rho_B \| \gamma_B)$, implying that there is an assisted protocol that helps Bob distilling more work. In particular, the states from which Bob can distil no work at all even in the assisted setting are simply products of the form $\Gamma_{AB} = \sigma_A \otimes \gamma_B$, from now on referred to as *quantum-thermal* (QT) states, the same states have been found in the conditional thermal operations setting [175].

In appendix A.3 we show that $W_a^{B|A}$ can be written as

$$W_a^{B|A}(\rho_{AB}) = \frac{1}{\beta}\left(S(\rho_B \| \gamma_B) + J^{\rightarrow}(\rho_{AB})\right), \tag{2.6}$$

where $J^{\rightarrow}(\rho_{AB})$ is the Henderson–Vedral [117] measure of classical correlations (with respect to measurements on Alice) defined as $J^{\rightarrow}(\rho_{AB}) := \max_{\{\Pi_{A,i}\}}\left(S(\rho_B) - \sum_i p_i S(\tilde{\rho}_{B,i})\right)$. The result in equation (2.6) clearly separates the quantity of work distillable by Bob with-/without the assistance of Alice. This is in agreement with a recent result in [163].

An important question to ask is whether this quantity of work changes if Alice is able to perform measurements over many copies of the shared initial state ρ_{AB}. In order to

answer this question we continue by defining the regularised work of assistance,

$$W_{a,\infty}^{B|A}(\rho_{AB}) := \lim_{n\to\infty} \frac{1}{n} W_a^{B|A}\left(\rho_{AB}^{\otimes n}\right). \tag{2.7}$$

In appendix A.4 we show that the above quantity indeed yields the best achievable rate of work distillation in the case where the only allowed communication is from Alice to Bob. Although the regularisation makes it hard to compute, the r.h.s. of (2.7) can nonetheless be related to a quantifier known as *distillable common randomness C_D*, introduced in [117] as

$$C_D^{\to}(\rho_{AB}) = \lim_{n\to\infty} \frac{1}{n} J^{\to}\left(\rho_{AB}^{\otimes n}\right), \tag{2.8}$$

and then interpreted operationally in [72]. The operational interpretation of C_D rests on protocols that extract from n independent copies of ρ_{AB} a total of C maximally correlated classical bits via R bits of noiseless classical communication between Alice and Bob with vanishing error. The quantity C_D is thus defined as the maximum net gain $(C - R)/n$ in the limit $n \to \infty$. For a discussion from the thermodynamical point of view, see [184].

Using the definition in equation (2.8) and the fact that the relative entropy is additive, we can therefore write the regularised work of assistance as,

$$W_{a,\infty}^{B|A}(\rho_{AB}) = \frac{1}{\beta}\left(S(\rho_B||\gamma_B) + C_D^{\to}(\rho_{AB})\right), \tag{2.9}$$

again clearly separating the quantity of distillable work with/without the assistance of Alice.

Upon defining the regularised version of $W_a^{B|A}(\rho_{AB})$ we should ask whether giving Alice the ability to perform global measurements over many copies of the shared state ρ_{AB} increases the average work that Bob can distil. In order to answer this question we employ two fundamental results from the field of quantum information. On the one hand, [143, Theorem 1] states that

$$E_f(\rho_{A'B}) + J^{\to}(\rho_{AB}) = S(\rho_B), \tag{2.10}$$

$$E_C(\rho_{A'B}) + C_D^{\to}(\rho_{AB}) = S(\rho_B), \tag{2.11}$$

provided that $\rho_{A'B}$ is the A-complement of ρ_{AB}, i.e. there exists a pure state extension $\rho_{AA'B}$ that satisfies $\text{Tr}_A\left[\rho_{AA'B}\right] = \rho_{A'B}$ and $\text{Tr}_{A'}\left[\rho_{AA'B}\right] = \rho_{AB}$. Here, $E_f(\rho_{AB})$ stands for the entanglement of formation [32], while the entanglement cost is given by $E_C(\rho_{AB}) = \lim_{n\to\infty} \frac{1}{n} E_f(\rho_{AB}^{\otimes n})$, and quantifies the amount of Bell states needed to form ρ_{AB} via LOCC protocols in the asymptotic limit of many copies [116].

Substituting equations (2.10)–(2.11) into equations (2.6)–(2.9) respectively allows us to write $W_a^{B|A}(\rho_{AB})$ and $W_{a,\infty}^{B|A}(\rho_{AB})$ in terms of these entanglement measures,

$$W_a^{B|A}(\rho_{AB}) = \frac{1}{\beta}\big(S(\rho_B||\gamma_B) + S(\rho_B) - E_f(\rho_{A'B})\big), \qquad (2.12)$$

$$W_{a,\infty}^{B|A}(\rho_{AB}) = \frac{1}{\beta}\big(S(\rho_B||\gamma_B) + S(\rho_B) - E_C(\rho_{A'B})\big). \qquad (2.13)$$

This allows us to take advantage of another fundamental result of quantum information, the non-additivity of $E_f(\rho_{AB})$ [114]. Therefore, despite the additivity of the (relative) von Neumann entropy we can state that the ability for Alice to perform global measurements can increase the amount of work Bob can distil, i.e. for some states ρ_{AB} it will happen that

$$W_a^{B|A}(\rho_{AB}) < W_{a,\infty}^{B|A}(\rho_{AB}). \qquad (2.14)$$

However, for many simple classes of states the above does not happen. For instance, in appendix A.5 we explicitly calculate $W_a^{B|A}$ for the relevant family of isotropic states in arbitrary dimension, and show its additivity over multiple copies.

2.4   Work of collaboration

Let us consider an arbitrary class of operations $\mathcal{O}$ on a thermodynamical system. We assume that $\mathcal{O}$ contains not only deterministic operations, but also so-called *quantum instruments*, i.e. collections $\{\Phi_i\}_i$ of completely positive maps such that $\sum_i \Phi_i$ is trace-preserving. Physically, the classical label i will record the outcomes of the quantum measurements that have been made throughout the process, while $\mathrm{Tr}\,\Phi_i(\rho)$ represents the probability of the outcome i occurring when the state ρ is processed. In a bipartite setting, we can construct the associated set $\mathcal{O}_c^{B|A}$ of collaborative operations by concatenating in any order: (1) instruments in $\mathcal{O}$ on B; (2) classical communication between Alice and Bob; (3) arbitrary quantum operations on A. We can now define the associated *work of collaboration* in analogy with Eq. (2.2) as

$$W_c^{B|A}(\rho_{AB}) := \sup\Big\{ RE : \qquad\qquad\qquad\qquad (2.15)$$
$$\lim_{n\to\infty}\inf_{\Lambda\in\mathcal{O}_c^{B|A}} \Big\|\Lambda\Big(\rho_{AB}^{\otimes n} \otimes |0\rangle\langle 0|_P^{\otimes[Rn]}\Big) - |1\rangle\langle 1|_P^{\otimes[Rn]}\Big\|_1 = 0\Big\},$$

where it is understood that the battery P pertains to Bob's system, and its Hamiltonian is again given by $H_P := E \, |1\rangle\langle 1|_P$, with E a free parameter.

By their very definition (2.1), TO are intrinsically deterministic. Therefore, in the collaborative setting there is no information Bob can send to Alice if he is restricted to TO, and the corresponding work of collaboration reduces to the regularised work of assistance as given in (2.7). To investigate the collaborative setting in greater detail it is thus indispensable to expand Bob's allowed operations to the wider class [82] of GP operations, that satisfy $\Lambda(\gamma_B) = \gamma_B$. This less restrictive framework crucially allows Bob to apply non-deterministic instruments $\{\Phi_i\}_i$, which are required to satisfy $\Phi_i(\gamma_B) \propto \gamma_B$ for all i. The outcome i can then be communicated to Alice via the classical communication channel.

From now on, we will therefore consider the work of collaboration (2.15) as defined for the collaborative set of operations $\mathrm{GP}_c^{B|A}$ corresponding to GP operations on Bob. It is clear that QT states of the form $\Gamma_{AB} = \sigma_A \otimes \gamma_B$, where σ_A is arbitrary, can be generated for free even in the TO's framework. Furthermore, it can be shown that these are all the states for which $W_c^{B|A}(\rho_{AB}) = 0$. This suggests the following definition of the *relative entropy of collaboration*,

$$W_r^{B|A}(\rho_{AB}) := \frac{1}{\beta} \min_{\sigma_A} S\left(\rho_{AB} \| \sigma_A \otimes \gamma_B\right), \qquad (2.16)$$

where the minimisation is taken over the set QT. In appendix A.6 we explicitly demonstrate monotonicity of this function under the set of allowed operations. We also prove in appendix A.6 that the minimisation in (2.16) can be explicitly solved so as to give

$$W_r^{B|A}(\rho_{AB}) = \frac{1}{\beta} S\left(\rho_{AB} \| \rho_A \otimes \gamma_B\right). \qquad (2.17)$$

Simple algebraic manipulations allow us to recast this as

$$W_r^{B|A}(\rho_{AB}) = \frac{1}{\beta}\left(S(\rho_B \| \gamma_B) + I(\rho_{AB})\right), \qquad (2.18)$$

where $I(\rho_{AB}) := S(\rho_A) + S(\rho_B) - S(\rho_{AB})$ is the mutual information quantifying total correlations between Alice and Bob.

2.5   Comparing measures of assistance

Equation (2.18) suggests that the mutual information quantifies the amount by which the collaboration between the parties increases Bob's distillable work.

In fact, we are able to demonstrate in appendix A.7 that $W_r^{B|A}$ provides an upper bound on the work of collaboration. We can also observe that since TO are a subset of GP operations, the work of collaboration is no smaller than the regularised work of assistance. This can also be deduced by comparing (2.9) with (2.18), and using the well-known fact that $C_D^{\rightarrow}(\rho_{AB}) \leq I(\rho_{AB})$ [71, 72]. Putting all together:

$$W_a^{B|A}(\rho_{AB}) \leq W_{a,\infty}^{B|A}(\rho_{AB}) \leq W_c^{B|A}(\rho_{AB}) \leq W_r^{B|A}(\rho_{AB}). \qquad (2.19)$$

Recall from (2.14) that there can be a strict inequality between the two leftmost quantities in the above chain of inequalities. Concerning the two rightmost ones, quite interestingly, we find that the gap $W_r^{B|A}(\rho_{AB}) - W_a^{B|A}(\rho_{AB})$ is explicitly described by the *quantum discord*, a measure of the quantumness of the correlations between Alice and Bob [183, 117]. Indeed, by comparing equations (2.6) and (2.18), we find

$$W_r^{B|A}(\rho_{AB}) - W_a^{B|A}(\rho_{AB}) = \frac{1}{\beta} \left(I(\rho_{AB}) - J^{\rightarrow}(\rho_{AB}) \right)$$
$$=: \frac{1}{\beta} D^{\rightarrow}(\rho_{AB}), \qquad (2.20)$$

where $D^{\rightarrow}(\rho_{AB})$ is the quantum discord, quantifying the share of correlations lost between Alice and Bob as a consequence of a minimally disturbing measurement on Alice's side. This result shows that the *work of collaboration* can exceed the *work of assistance* by an amount bounded from above by the shared quantum correlations, measured by the discord $D^{\rightarrow}(\rho_{AB})$. We note that recent works [86, 163] has suggested a protocol for explicitly distilling the work locked in the quantum discord, however the operations considered lie outside those in TO. Other interpretations for the quantum discord in thermodynamical and related contexts have also been explored in the literature [256, 184, 155, 6].

It is particularly instructive to analyze all the quantities appearing in equation (2.19) for the relevant case where Alice holds a purification of Bob's state, i.e. $\rho_{AB} = \phi_{AB} = |\phi\rangle\langle\phi|_{AB}$. On the one hand, for a pure state ϕ_{AB} it is known [117, 72] that the Henderson–Vedral measure and distillable common randomness coincide with the local entropy of each subsystem, i.e. $J^{\rightarrow}(\phi_{AB}) = C_D^{\rightarrow}(\phi_{AB}) = S(\phi_B)$. Hence,

$$W_a^{B|A}(\phi_{AB}) = W_{a,\infty}^{B|A}(\phi_{AB}) = \frac{1}{\beta} \left(S(\rho_B || \gamma_B) + S(\rho_B) \right), \qquad (2.21)$$

implying that for an initial pure state the ability for Alice to perform global measurements over many copies gives no advantage in Bob distilling work. On the other hand, it is also elementary to verify that

$$W_r^{B|A}(\phi_{AB}) = \frac{1}{\beta}\left(S(\rho_B||\gamma_B) + 2\,S(\rho_B)\right).$$

(2.22)

Therefore by comparing equations (2.21) and (2.22) it is seen that for an initial pure state we demonstrate that relaxing the local operations from TO to GP map might allow Bob to distil a bound quantity of work equal to the local entropy.

2.6   Discussion

In this chapter we have fully characterised the task of assisted work distillation in the asymptotic scenario of quantum thermodynamics, addressing questions left open in [58, 184]. In particular we have introduced two relevant quantities of interest, the *work of assistance* and the *work of collaboration.* These quantities allowed us to investigate the possible advantage of local GP operations over TO and global measurements on a system; in particular, how GP operations may allow Bob to locally distil the work bound within the quantum correlations of the initial shared state.

Although it was shown that GP operations can provide an increase in distillable work, the explicit relationship between the *work of assistance* and the *work of collaboration* requires further investigation, as for the latter quantity only an upper bound was derived here. We further stress that our results only hold in the asymptotic limit. It would be interesting to investigate assisted work distillation in the *single-shot regime*, to determine the role correlations play in work fluctuations. This could prove useful for near-term technological applications, such as microscopic heat engines [45] or may even have applications in algorithmic cooling [186].

The present analysis adds to the literature on assisted distillation of different quantum resources [75, 124, 70, 58, 198, 217]. In particular, Refs. [58, 198] studied the distillation of quantum coherence [216], rather than work, from Bob's system with the assistance of Alice. In that setting, Bob is limited to incoherent operations [24] while Alice can perform arbitrary local quantum operations, and the two parties can communicate classically. We can draw a comparison between the two settings, by noting that the additional quantity of resource that can be distilled from Bob's system thanks to Alice's assistance amounts to the entropy of Bob's reduced state in the case of coherence [58] and to the classical correlations shared between Alice and Bob in the case of work [Eq. (2.6)]. We can further

observe how the hierarchy presented in (2.19) for assisted work distillation is analogous to the one derived in [58] for assisted coherence distillation, but the key role of quantum discord in bounding the gap between work of assistance and work of collaboration is only revealed in this chapter by comparing the power of different classes of local operations for Alice (TO versus GP). It would be meaningful to revisit the assisted coherence distillation framework by imposing additional physical constraints on Alice's operations, e.g. by adopting strictly incoherent operations [242] or TO, and hence exploiting the methods developed in this chapter for the characterisation of other quantum resources.

Our findings also have implications for the understanding of the Szilard engine [222]. The latter is a simple physical model which demonstrates how information may be exploited in order to extract physical work. The relevance of this model was then understood in the context of information processing by Landauer [151]. Many recent works have discussed the application of a Szilard engine in quantum thermodynamics [141, 167, 63, 187, 255, 196, 201], deriving bounds for work extraction that are related to (2.18) [187, 255, 196, 201] in a setting where a second party, historically entitled Maxwell's Demon, is in possession of a state correlated to the thermodynamic system. The converse setting, where correlations can be formed from initially uncorrelated states using thermal operations has also been studied [176].

The results presented here provide further links between the fields of quantum information and thermodynamics. In particular, how highly studied measures of information provide us with an insight into the thermodynamics of correlations. These results both contribute to our knowledge of the fundamental nature of thermodynamics but also may become essential for the thermodynamic control of a quantum computer.

Chapter 3

Quantum coherence fluctuation relations

3.1 Introduction

As discussed in chapter 1 the development of a resource theory of quantum coherence mirrors the early motivation behind the theoretical investigations of classical thermodynamics, where optimal procedures were derived for distilling work from a thermal machine [50]. These have been superseded by the fields of stochastic and quantum thermodynamics [145, 103], most notably by the seminal fluctuation theorems due to Jarzynski [132] and Crooks [64], see section 1.4, which consider the amount of extractable work as a quantity that can fluctuate during a thermodynamic process, and hence characterise fundamental limitations on the associated work distribution.

Recent work has formalised a connection between the algebraic theory of majorisation and the emergence of fluctuation theorems [8, 9]. This has been highlighted not only in thermodynamics [8], where so-called thermo-majorisation provides necessary and sufficient conditions for state transformations under thermal operations within the resource theory of athermality [127, 105], but also in the context of entanglement theory [9], where pure state transformations under local operations and classical communication (LOCC) are once again determined by majorisation relations [179]. These observations raise the prospect that other resources, *in primis* coherence, may also be allowed to fluctuate and give rise to a distribution regulated by fluctuation theorems while implementing the conversion of quantum states under the corresponding set of free operations.

In this chapter we establish fluctuation relations for the manipulation of quantum co-

herence under incoherent or strictly incoherent operations [24, 242, 248], see section 1.5.2, that is, another instance where majorisation theory provides necessary and sufficient conditions for pure state transformations $|\psi\rangle_A \to |\phi\rangle_A$ in a quantum system A [242, 55, 254]. In order to do this, an ancillary device that stores and supplements coherence is necessary, introduced here as a *coherence battery B*. This battery, initialised in a state $|\lambda\rangle_B$, is used as an approximate catalyst to mediate the pure state transformation $|\Psi\rangle_{AB} \to |\Phi\rangle_{AB}$ as

$$|\Psi\rangle_{AB} = |\psi\rangle_A \otimes |\lambda\rangle_B \to |\Phi\rangle_{AB} \approx |\phi\rangle_A \otimes |\lambda\rangle_B, \qquad (3.1)$$

where the approximation becomes exact and the transformation reversible in the limit of an ideal battery, as discussed later in Section 3.3. This establishes a resource-theoretic framework for coherence manipulation under battery assisted incoherent operations (BIO) or battery assisted strictly incoherent operations (BSIO), collectively referred to as B(S)IO. Necessary and sufficient conditions are derived for the battery to gain or lose a quantity of coherence w probabilistically. This gives rise to a coherence distribution $P(w)$ following the transformation. From this distribution, four theorems characterising fundamental limitations on the manipulation of fluctuating coherence are then derived.

Result 1— A second law of coherence is derived, which governs the amount of coherence extractable from the battery during the transformation (in the same sense that the second law of thermodynamics governs the amount of extractable work during a transformation). If coherence is allowed to fluctuate, we find that the average extractable coherence is bounded by the difference in relative entropy between initial and final states of the system. This complements the fact that the relative entropy of coherence [24] yields the exact distillable coherence under incoherent operations within the standard resource theory of coherence [242].

Result 2— A third law of coherence is derived, which demonstrates that the limiting factor for extracting fluctuating coherence is the diagonal rank of the density matrix. This again mirrors the standard result, namely the rank of the diagonal part of pure state density matrices cannot increase under incoherent operations [242].

Result 3— An analogue to Jarzynski's relation [132] is derived, which applies when the final state of the system is maximally coherent. This shows the nature of fluctuating coherence and, in conjunction with the second and third laws of coherence, demonstrates strong bounds on extractable coherence during the transformation.

Result 4— By comparing forward and reverse transformations, an analogue of Crooks' relation [64] is found, which applies when the final states of both transformations are

maximally coherent. It implies that extracting w units of coherence from the battery in the forward protocol is exponentially suppressed with respect to extracting $-w$ units in the reverse protocol, showing an inherent irreversibility in coherence manipulation.

This chapter is organised as follows. Section 3.2 presents the conditions for state transformations under (strictly) incoherent operations defined by the majorisation criteria. Section 3.3 characterises the coherence battery employed to mediate pure state transformations. Section 3.4 describes necessary and sufficient conditions for battery assisted state transformations, detailing the protocol that gives rise to the fluctuating coherence distribution. Section 3.5 presents the aforementioned four results governing the fluctuation relations for coherence. Section 3.6 contains a summary and discussion of our results. Within the Appendix are proofs of the conditions for battery assisted transformations and the derivation of the reverse protocol necessary for the coherence analogue of Crooks' theorem.

Throughout this chapter, the density matrix of a pure state $|\psi\rangle_A$ will be denoted by ψ_A, the subscript indicating the subsystem to which the state belongs (usually A for the principal system, and B for the battery). The Hilbert spaces of system A and battery B will be denoted by $\mathcal{H}_A$ and $\mathcal{H}_B$, and the corresponding set of density matrices by $\mathcal{D}(\mathcal{H}_A)$ and $\mathcal{D}(\mathcal{H}_B)$, respectively. Occasionally subsystem labels will be omitted when clear from the context.

3.2   State transformation

In this chapter we consider a pure to pure state transformation

$$|\Psi\rangle \to |\Phi\rangle \, , \tag{3.2}$$

It is known that such a transformation is possible by means of general deterministic SIO or IO, that is, $\exists \Lambda \in$ (S)IO such that $\Phi = \Lambda(\Psi)$, if and only if [76, 77, 242, 55, 54, 56, 254]

$$\Delta(\Psi) \prec \Delta(\Phi) \, , \tag{3.3}$$

that is, equation [?] shows that $\Delta(\Psi)$ is majorised by $\Delta(\Phi)$. Explicitly, the necessary and sufficient condition for this majorisation relation to hold is [37, 181]

$$\Delta(\Psi) = \sum_m r_m \Xi_m \Delta(\Phi) \Xi_m^\dagger \, , \tag{3.4}$$

where $r_m \geq 0$, $\sum_m r_m = 1$, and Ξ_m are permutation matrices. Defining this in terms of a completely positive trace preserving unital map E acting on any operator X,

$$\mathrm{E}(X) := \sum_m r_m \Xi_m X \Xi_m^\dagger, \tag{3.5}$$

we then have that, as depicted in Figure 3.1(a), the pure state transformation in Eq. (3.2) can be implemented by (S)IO if and only if there exists a unital map E of the form (3.5) such that

$$\Delta(\Psi) = \mathrm{E}(\Delta(\Phi)). \tag{3.6}$$

This is equivalent to the existence of a bistochastic matrix mapping the (nonzero) diagonal coefficients of Φ to those of Ψ [37].
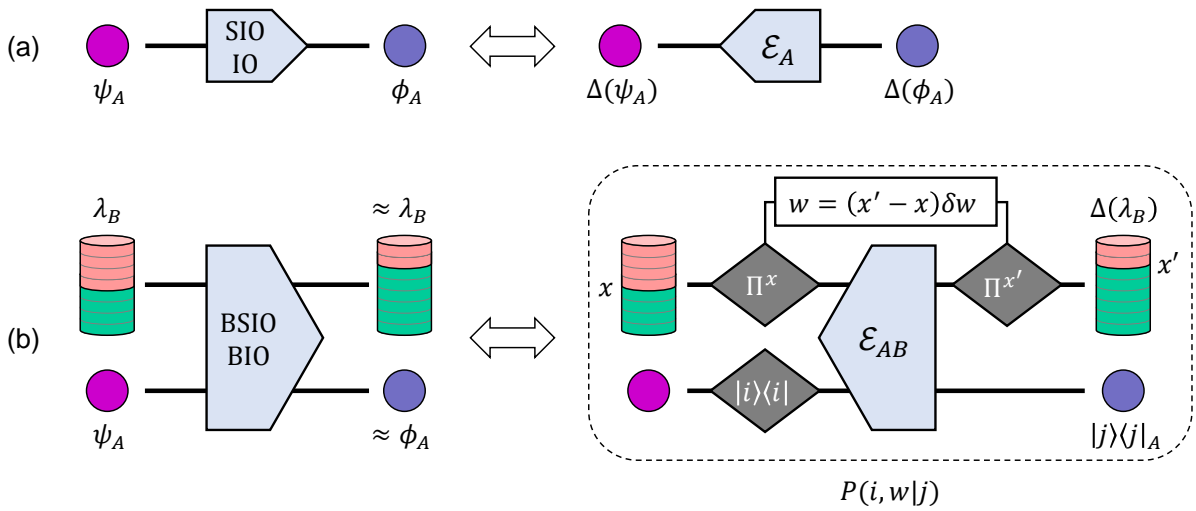
Figure 3.1: **(a)** In the standard resource theory of quantum coherence [216], the pure to pure state transformation $|\psi\rangle_A \to |\phi\rangle_A$ on the system A (left) can be implemented by SIO or IO if and only if there exists a unital map E_A of the form (3.5) that maps the diagonal component of the final state $\Delta(\phi_A)$ into that of the initial state $\Delta(\psi_A)$ (right). **(b)** In the battery assisted framework considered here, the pure to pure state transformation $|\psi\rangle_A \otimes |\lambda\rangle_B \to |\Phi\rangle_{AB} \approx |\phi\rangle_A \otimes |\lambda\rangle_B$ (left) can be implemented by BSIO or BIO — that is, by SIO or IO on the system A and the battery B, accompanied by a change in coherence of the battery by an amount w with probability $P(w)$ — if and only if there exists a conditional probability distribution $P(i, w|j)$ that satisfies the three conditions given in Eqs. (3.25)–(3.27). Such a distribution can be constructed from the statistics of the protocol illustrated in the dashed box (right), described in Section 3.4. This framework allows us to investigate fluctuation relations for quantum coherence, analogous to those for work in thermodynamics, as presented in Section 3.5.

3.3   Coherence battery

We consider a system A on which we aim to perform the pure to pure state transformation $|\psi\rangle_A \to |\phi\rangle_A$, supplemented by a battery B, so that the composite state transformation can be written overall as in Eq. (3.1). The battery is initialised in a state $|\lambda\rangle_B$ that can be defined in general as a superposition of coherence eigenstates $|c_x\rangle$,

$$|\lambda\rangle_B := \sum_{x=0}^{n} \sqrt{\alpha_x} |c_x\rangle_B \,, \tag{3.7}$$

with $\alpha_x \geq 0$, $\sum_x \alpha_x = 1$. Here by coherence eigenstates we mean states $|c_x\rangle$ with a well defined amount of coherence, as quantified by the relative entropy C_{rel}. In particular, in analogy to the case of the entanglement battery studied in [9], we can write each $|c_x\rangle$ as the tensor product of two types of states, namely x copies of a state $|\Upsilon_u^+\rangle$ with higher coherence (i.e., a charged state) and $n - x$ copies of a state $|\Upsilon_u^-\rangle$ with lower coherence (i.e., a discharged state). Precisely,

$$|c_x\rangle := \underbrace{|\Upsilon_u^+\rangle \otimes .... \otimes |\Upsilon_u^+\rangle}_{x} \otimes \underbrace{|\Upsilon_u^-\rangle \otimes ..... \otimes |\Upsilon_u^-\rangle}_{n-x}, \tag{3.8}$$

for chosen integers n and u and for all integers $x \in \{0, ..., n\}$, with

$$|\Upsilon_u^+\rangle := \frac{1}{\sqrt{u}} \sum_{i=1}^{u} |i\rangle \,, \tag{3.9}$$

$$|\Upsilon_u^-\rangle := \frac{1}{\sqrt{u-1}} \sum_{i=u+1}^{2u-1} |i\rangle \,. \tag{3.10}$$

From Eq. (1.54), we see that the relative entropy of coherence of the charged and discharged states is given respectively by

$$C_{\text{rel}}(\Upsilon_u^+) = \ln(u), \tag{3.11}$$

$$C_{\text{rel}}(\Upsilon_u^-) = \ln(u-1). \tag{3.12}$$

Note that the states $|\Upsilon_u^+\rangle$ and $|\Upsilon_u^-\rangle$ are equivalent to maximally coherent states of dimension u and $u - 1$, respectively. The coherence of the state (3.8) is then given by

$$C_{\text{rel}}(c_x) = \ln\left(u^x(u-1)^{n-x}\right) \,. \tag{3.13}$$

A measurement of the 'position' of the battery, or more properly, of its level of coherence as specified by the index x, can be obtained by defining a set of orthogonal projectors Π^x

as

$$\Pi^x := u^x(u-1)^{n-x}\chi_x \, , \tag{3.14}$$

where χ_x is the incoherent state corresponding to the diagonal part of the state (3.8),

$$\chi_x := \Delta(|c_x\rangle\langle c_x|) = \Big(\frac{1}{u}\sum_{i=1}^{u}|i\rangle\langle i|\Big)^{\otimes x} \otimes \Big(\frac{1}{u-1}\sum_{i=u+1}^{2u-1}|i\rangle\langle i|\Big)^{\otimes(n-x)} \, . \tag{3.15}$$

By construction, $\sum_x \Pi^x =: \mathbb{B}_B$, which is the projector on the diagonal support of the battery; in other words, the projectors $\{\Pi^x\}$ give a resolution of the identity on the subspace $\mathcal{B} \subset \mathcal{D}(\mathcal{H}_B)$ of the state space of the battery, spanned by $\mathrm{supp}\big(\Delta(\lambda_B)\big) = \bigoplus_{x=0}^{n} \mathrm{supp}(\chi_x)$.

Similarly to a conventional energy storage device, the battery B will act as a coherence supplier that can receive/transfer coherence from/to the system A, by changing the ratio of the states (3.9) and (3.10). In fact, the discharging process $|\Upsilon_u^+\rangle \to |\Upsilon_u^-\rangle$ corresponds to decreasing x by one and hence diminishes the coherence in the battery by a *quid* δw,

$$\delta w := \ln\Big(\frac{u}{u-1}\Big) \, . \tag{3.16}$$

This can be seen as extracting one unit of coherence from the battery.

In general, the role of the battery is to mediate the state transformation (3.1) by exchanging an amount w of coherence with the system. We may choose the parameter u large enough, corresponding to a level spacing $\delta w \approx 1/u$ in the battery fine enough, so that any change w in the coherence of the battery can be taken approximately to be a multiple of δw. Ideally, we would like the battery to be reusable in order to assist subsequent state transformations. Furthermore, we would like the battery to serve the purpose of overcoming the limitations in conventional (unassisted) state transformations under (S)IO on the system, going beyond the conditions of Section 3.2. Therefore, there are three constraints that an ideal battery should adhere to:

1. In order for the final state $|\phi\rangle_A$ to be pure, the system should be virtually uncorrelated with the battery, $|\Phi\rangle_{AB} \approx |\phi\rangle_A \otimes |\lambda\rangle_B$.

2. The only allowed action on the battery should be the raising and lowering of w units of coherence, by the unitary operator Γ^w defined as

$$\Gamma^w|c_x\rangle_B = |c_{x+w}\rangle_B \, , \tag{3.17}$$

with $x + w$ assumed modulo $n + 1$, and x and n assumed large enough to avoid hitting the bottom or top levels of the battery.

3. The state of the battery $|\lambda\rangle_B$ should allow for approximately implementing all reversible pure to pure state transformations $|\psi\rangle_A \to |\phi\rangle_A$ by (S)IO.

The first constraint is fulfilled provided the chosen battery state $|\lambda\rangle_B$ is a superposition over sufficiently many eigenstates $|c_x\rangle$, that is, provided the size of the battery, determined by the parameter n, is chosen large enough. The last two constraints are stronger and force the state of the battery $|\lambda\rangle_B$ to be close to a uniform superposition of coherence eigenstates $|c_x\rangle$. To see this, note that the second constraint imposes that the final state of system and battery has to be of the form

$$|\Phi\rangle_{AB} = \sum_{w:|w|\leq w_{\max}} |\phi_w\rangle_A \otimes \Gamma^w |\lambda\rangle_B \,, \tag{3.18}$$

for some $w_{\max} > 0$, while the third constraint imposes that, for all reversible transformations (3.1) implemented by (S)IO, the final state is ϵ-close to the target one,

$$||\Phi_{AB} - \phi_A \otimes \lambda_B||_1 \leq \epsilon \,, \tag{3.19}$$

and with identical diagonal marginal on the system,

$$\Delta(\mathrm{Tr}_B \Phi_{AB}) = \Delta(\phi_A) \,. \tag{3.20}$$

As the conditions above have to hold for *all* reversible state transformations, we can analyse the specific one where the initial and final states of the system are, respectively,

$$|\psi\rangle_A = \frac{1}{\sqrt{2}}(|0\rangle_A + |\Upsilon_u^+\rangle_A) \,, \tag{3.21}$$

$$|\phi\rangle_A = \frac{1}{\sqrt{2}}(|0\rangle_A + |\Upsilon_u^-\rangle_A) \,. \tag{3.22}$$

Considering now system and battery initialised in the state $|\Psi\rangle_{AB} = |\psi\rangle_A \otimes |\lambda\rangle_B$, and noting that the transformation $|\Psi\rangle_{AB} \to |\Phi\rangle_{AB}$ can be implemented reversibly under SIO [242] or IO [55, 254] if and only if the nonzero diagonal coefficients of the initial and final states are identical, one can show that the only final state of system and battery that fulfils this requirement while complying with Eqs. (3.18) and (3.20) is

$$|\Phi\rangle_{AB} = \frac{1}{\sqrt{2}}(|0\rangle_A \otimes |\lambda\rangle_B + |\Upsilon_u^-\rangle_A \otimes \Gamma^{\delta w} |\lambda\rangle_B). \tag{3.23}$$

The proof follows closely the one reported in [9] for LOCC transformations in entangle-

ment theory, with diagonal coefficients here playing the same role as Schmidt coefficients there. Finally invoking Eq. (3.19) and applying further algebra [9], one finds

$$\sum_x |\alpha_x - \alpha_{x+y}| \leq |y|\sqrt{8\epsilon}\,, \quad \forall |y| \leq w_{\max}/\delta w\,. \tag{3.24}$$

This means that, in order for the battery to serve as an approximate catalyst to implement all reversible (S)IO pure to pure state transformations, including the specific instance just discussed, the set of coefficients α_x in its initial superposition state $|\lambda\rangle_B$ of the form (3.7) must be close to a uniform distribution, as formalised by Eq. (3.24).

3.4   Battery assisted state transformations and coherence distribution protocol

We are now ready to investigate necessary and sufficient conditions for the transformation between initial and final states $|\psi\rangle_A$ and $|\phi\rangle_A$ of the system A, with diagonal components $\Delta(\psi_A) = \sum_i p_i |i\rangle\langle i|_A$ and $\Delta(\phi_A) = \sum_j q_j |j\rangle\langle j|_A$, mediated by a change in coherence of an amount w with probability distribution $P(w)$ in the battery B, initially prepared in the state $|\lambda\rangle_B$ of Eq. (3.7). Here the distribution $P(w)$ is associated to a two-stage measurement of the battery with the projectors (3.14) before and after the transformation, that is, $P(w)$ is the probability of finding the battery in the final state $|c_{x+w}\rangle_B$, given that it was found initially in the state $|c_x\rangle_B$.

The main result of this Section, which mirrors the analogous one recently reported for entanglement theory [9], is illustrated in Figure 3.1(b) and can be enunciated as follows.

Theorem 3 (Necessary and sufficient conditions for B(S)IO transformations)**.** *The transformation of Eq. (3.1) can be implemented by means of SIO or IO on the system and the battery while extracting a coherence distribution $P(w)$ — that is, by battery assisted (S)IO or, in short, B(S)IO — if and only if there exists a conditional probability distribution $P(i,w|j)$, with marginals $P(i) = p_i$, $P(j) = q_j$, and $P(w)$, which fulfils the following three conditions:*

$$\text{Condition 1: } \sum_{i,w} P(i,w|j) = 1,\ \ \forall j\,, \tag{3.25}$$

$$\text{Condition 2: } \sum_{j,w} P(i,w|j)e^w = 1,\ \ \forall i\,, \tag{3.26}$$

$$\text{Condition 3: } \sum_{j,w} P(i,w|j)q_j = p_i,\ \ \forall i\,. \tag{3.27}$$

Physically, Condition 1 expresses the normalisation of the conditional probability distribution $P(i, w|j)$, Condition 2 regulates the fluctuations of w units of coherence in the battery with probability $P(w)$, while Condition 3 formalises the requirement that the marginals $P(i)$ and $P(j)$ of the joint probability distribution $P(i, j, w) = P(i, w|j)P(j)$ reproduce the diagonal components p_i and q_j of the initial and final states of the system, respectively.

Proof. The proof of the Theorem consists of two directions. For the "if" part, we need to show that, given a conditional probability distribution $P(i, w|j)$ obeying Conditions 1–3, sequences of B(S)IO protocols $\Lambda_{AB}^{(N)}$ and states $\left|\Psi^{(N)}\right\rangle_{AB}$ and $\left|\Phi^{(N)}\right\rangle_{AB}$ exist, such that $\Phi_{AB}^{(N)} = \Lambda_{AB}^{(N)}(\Psi_{AB}^{(N)})$, with $\lim_{N\to\infty} \left|\Psi^{(N)}\right\rangle_{AB} = |\psi\rangle_A \otimes |\lambda\rangle_B$ and $\lim_{N\to\infty} \left|\Phi^{(N)}\right\rangle_{AB} = |\phi\rangle_A \otimes |\lambda\rangle_B$. As recalled in Section 3.2, this is equivalent to showing the existence of a sequence of bistochastic matrices $G^{(N)}$ mapping the (nonzero) diagonal coefficients of $\Phi_{AB}^{(N)}$ to those of $\Psi_{AB}^{(N)}$. Such a derivation is rather technical and hence deferred to B.1.1.

For the "only if" part, let us assume that a B(S)IO transformation (3.1) is possible, that is, there exists a unital map E_{AB} of the form (3.5) such that the (nonzero) diagonal components of the initial and final states satisfy

$$\Delta(\Psi_{AB}) = \mathrm{E}_{AB}(\Delta(\Phi_{AB})) \, . \tag{3.28}$$

We then need to prove that a conditional probability distribution $P(i, w|j)$ fulfilling the above three conditions exists. It turns out one can explicitly construct such a probability distribution from the following five step protocol, also schematically represented in Figure 3.1(b):

1. Prepare the incoherent state $|j\rangle\langle j|_A \otimes \Delta(\lambda_B)$;

2. Measure the battery with the projector $\Pi_B^{x'}$ from (3.14);

3. Transform the resulting state of system and battery with the unital map $\mathcal{E}_{AB}$ of Eq. (3.28);

4. Measure the system with the projector $|i\rangle\langle i|_A$ and the battery with $\Pi_B^x \equiv \Pi_B^{x' - \frac{w}{\delta w}}$;

5. Record the variable $w = (x' - x)\delta w$, discarding x and x'.

The protocol above, in which $\frac{w}{\delta w}$ describes the amount w of extracted coherence as a multiple of the unit δw defined in (3.16), gives rise to the probability distribution

$$P(i, w|j) = \sum_{x'} \mathrm{Tr}[(|i\rangle\langle i|_A \otimes \Pi_B^{x' - \frac{w}{\delta w}})\mathcal{E}_{AB}(|j\rangle\langle j|_A \otimes \Pi_B^{x'} \Delta(\lambda_B)\Pi_B^{x'})] \, , \tag{3.29}$$

which, using Eq. (3.14), can also be rewritten as

$$P(i,w|j) = \sum_{x'} \alpha_{x'} \text{Tr}[(|i\rangle\langle i|_A \otimes \Pi_B^{x'-\frac{w}{\delta w}}) \mathcal{E}_{AB}(|j\rangle\langle j|_A \otimes \chi_{x'B})]. \qquad (3.30)$$

The proof that $P(i,w|j)$ satisfies Conditions 1–3 is reported in B.1.2. $\square$

3.5   Fluctuation theorems from coherence distribution

We have shown that the amount w of fluctuating coherence exchanged between battery and system when mediating a pure to pure state transformation $|\psi\rangle_A \to |\phi\rangle_A$ gives rise to a conditional probability distribution (3.30). In analogy to the derivation of the fluctuation theorems from a conditional work probability distribution in thermodynamics [64], several coherence fluctuation theorems can now be obtained. In this section, we present the mathematical derivation of the four main results anticipated in section 3.1, accompanying each of them with relevant physical remarks and comparisons with the corresponding thermodynamic laws.

3.5.1   Second law of coherence

The following is for an initial state $|\psi\rangle_A$ with diagonal coefficients p_i and a target state $|\phi\rangle_A$ with diagonal coefficients q_j.

Starting with Condition 2, multiplying Eq. (3.26) by p_i, rewriting the conditional probability distribution as $P(i,w|j) = P(i,j,w)/q_j$, and summing over i gives

$$\sum_{i,j,w} P(i,j,w) \frac{p_i}{q_j} e^w = 1, \qquad (3.31)$$

where we have used Condition 3. Now using $\frac{p_i}{q_j} = e^{\ln\left(\frac{p_i}{q_j}\right)}$ to move the probabilities into the exponent,

$$\sum_{i,j,w} P(i,j,w) e^{w - \ln q_j + \ln p_i} = 1, \qquad (3.32)$$

and writing in bracket form, we get

$$\left\langle e^{w - \ln q_j + \ln p_i} \right\rangle = 1 \,. \tag{3.33}$$

This describes the distribution of fluctuating coherence w that can be extracted during our pure state transformation. By expanding to first order, and using Eq. (1.54), we find

$$\langle w \rangle \leq C_{\mathrm{rel}}(\psi_A) - C_{\mathrm{rel}}(\phi_A) \,. \tag{3.34}$$

This shows that the *average* coherence extractable from the battery to mediate B(S)IO state transformations is bounded by the difference in relative entropy of coherence between the initial and final states of the system. This is in contrast to the standard operational setting in the resource theory of quantum coherence, in which the relative entropy of coherence (scaled by a factor $\ln(2)$ in our notation) quantifies the exact distillable coherence under IO [242].

We can also see that Eq. (3.34) is formally analogous to the traditional second law of thermodynamics, $\langle W \rangle \leq F(\rho) - F(\sigma)$, which states that during the state transformation $\rho \rightarrow \sigma$ the average work W required is less than or equal to the difference in free energies $F(\rho) = \langle H \rangle - TS(\rho)$. As Eq. (3.34) is the first order expansion, this is just the average result. Higher order Taylor expansions of Eq. (3.33) lead to all the moments of the coherence distribution $P(w)$ that can be obtained during the transformation.

3.5.2 Third law of coherence

The following is for the transformation $|\psi\rangle_A \rightarrow |\phi\rangle_A$, where p_{min} and q_{min} are the smallest nonzero diagonal coefficients of the initial and final state, respectively.

Starting with Condition 3 on the probability distribution (3.27),

$$\sum_{j,w} P(i_0, w | j) q_j = p_{\min} \,, \tag{3.35}$$

where i_0 is the index corresponding to the smallest diagonal coefficient of the state, $p_{i_0} \equiv p_{\min}$, we can write the inequality $P(i_0, w | j) q_j \leq p_{\min} \; \forall j$, that can be substituted into Eq. (3.26), to get

$$1 \leq \sum_{j,w} \frac{p_{\min}}{q_j} e^w \leq \sum_{j,w} \frac{p_{\min}}{q_{\min}} e^w \,, \tag{3.36}$$

where we further used the fact that $1/q_j \leq 1/q_{\min} \; \forall j$ by construction. Summing the rightmost term over j, the diagonal rank d' of the final state is obtained, so that we can write

$$\sum_w e^w \geq \frac{q_{\min}}{d' p_{\min}} \,. \tag{3.37}$$

This can be interpreted as the third law for fluctuating coherence. It is well known that the majorisation criterion (3.3) for state transformations in the resource theory of coherence implies the following statement, namely that the rank of the diagonal part of pure states cannot increase under (S)IO [242]. Here, we find that the amount of fluctuating coherence w required to increase the diagonal rank (i.e., to send $p_{\min} \to 0$) under B(S)IO must diverge. Therefore such an operation is forbidden as it would require a battery of infinite size. The analogous result in thermodynamics is that decreasing the rank of a density matrix requires infinite resources, which can be regarded as a general statement of the third law [166].

3.5.3   Jarzynski's relation for coherence

The following is for the transformation $|\psi\rangle_A \to |\phi\rangle_A$, where the final state is a maximally coherent state of dimension d'.

Starting again from Condition 2, and using the fact that $q_j = 1/d' \; \forall j$ for a maximally coherent final state, we can multiply both sides of Eq. (3.26) by $1/d'$ and obtain

$$\sum_{j,w} P(i,j,w) e^w = \frac{1}{d'} \,. \tag{3.38}$$

Now summing over the index i gives

$$\sum_{i,j,w} P(i,j,w) e^w = \sum_i \frac{1}{d'} = \frac{d}{d'} \,, \tag{3.39}$$

which can again be written in bracket form,

$$\langle e^w \rangle = \frac{d}{d'} \,, \tag{3.40}$$

where d and d' denote the diagonal rank of the initial and final states, respectively.

In statistical mechanics, Jarzynski's relation [132] $\langle e^{\beta W} \rangle = \frac{Z'}{Z}$ describes an initial ther-

mal state which is driven out of equilibrium to a final state with a different Hamiltonian H', where Z and Z' are the initial and final partition functions, respectively. Jarzynski's relation implies that, when trying to extract work from a thermal bath, the probability of success decreases exponentially with the amount of work W being extracted. Equation (3.40) equivalently says that attempting to extract more coherence than the average for an initial state of dimension d results in the dimension of the maximally coherent final state, d', decreasing. As known by the analogous third law of coherence in Eq. (3.37), decreasing the dimension of the final state is limited by the maximum fluctuating coherence w and thus by the size of the battery. Explicitly, from Eq. (3.40) it follows that $P\big(w \geq \ln\big(\tfrac{d}{d'}\big) + r\big) \leq e^{-r}$.

The comparison between Eq. (3.40) and Jarzynski's equation also highlights an analogy between d/d' and Z'/Z. Using the relation from statistical mechanics $F = -1/\beta \ln(Z)$, where F is the free energy of the state, the ratio $Z'/Z = e^{\beta(F-F')}$ is describing the exponential of the extractable work from the state under thermal operations. Similarly for d/d', using the relation $C_{\mathrm{rel}} = \ln d$ for a pure (maximally coherent) state, the ratio $d/d' = e^{C'_{\mathrm{rel}} - C_{\mathrm{rel}}}$ is expressing the exponential of the extractable coherence from the state under B(S)IO. Note also that the fluctuation relation in Eq. (3.40) holds for a whole family of (not necessarily maximally coherent) initial states with the same d; this is mirrored by the redefinition of free energy from a single average value to a family of free energies with the same fluctuating behaviour [8], giving rise to the 'many second laws' of quantum thermodynamics [44].

3.5.4 Crooks' relation for coherence

The following is for the transformation $|\psi\rangle_A \to |\phi\rangle_A$ and its reverse $|\psi'\rangle_A \to |\phi'\rangle_A$, where the final states of the forward and reverse transformations are maximally coherent states of dimension d' and d respectively.

Crooks' theorem of statistical mechanics [64] relates the forward and reverse probabilities of extracting a quantity of work W during a non-equilibrium transformation between two thermal states,

$$\frac{P(W)}{P^{\mathrm{rev}}(-W)} = e^{-\beta W} \frac{Z'}{Z} \,. \tag{3.41}$$

Crooks' equation shows that the forward protocol is exponentially suppressed in comparison to its reverse. This is a quantitative description of the emergent irreversibility of thermodynamics.

In order to find the coherence analogue of this relation, a reverse transformation protocol is derived in B.2, where an integer quantity of coherence $-\frac{w}{\delta w}$ is extracted. This results in forward and reverse coherence distributions of the form

$$P(w) = \sum_{i,j} P(i, w|j) \frac{1}{d'}, \tag{3.42}$$

$$P^{\mathrm{rev}}(-w) = \sum_{i,j} P^{\mathrm{rev}}(j, -w|i) \frac{1}{d}. \tag{3.43}$$

It is shown in B.2 that these distributions obey the relation

$$\frac{P(w)}{P^{\mathrm{rev}}(-w)} = e^{-w} \frac{d}{d'}. \tag{3.44}$$

In analogy to Crooks' theorem it can be seen that extracting w units of coherence in the forward protocol is exponentially less likely than extracting $-w$ in the reverse protocol. One could attempt to increase the preference of the forward protocol $P(w)$ by decreasing the diagonal rank d' of its final state, however according to the equivalent third law of coherence in Eq. (3.37) this is exponentially difficult in its own right. It has therefore been shown that there is an inherent irreversibility in the manipulation of coherence within the B(S)IO framework.

3.6 Discussion

In this chapter we have established fluctuation relations for the manipulation of quantum coherence, in the context of pure to pure state transformations via (strictly) incoherent operations, assisted by a coherence battery which can exchange coherence with the system probabilistically. We hope that this work motivates a further reconsideration of coherence in quantum mechanics, from a useful albeit static resource that may be invested to convert quantum states and realise useful technological applications, to a quantity that more generally can fluctuate during transformations and may enable otherwise impossible applications. This has been accomplished here, similarly to the way in which work has been redefined in quantum thermodynamics [8] and more recently entanglement has been investigated as a fluctuating quantity in quantum information theory [9]. It is surprising that by forming parallels between coherence, entanglement, and work distribution protocols, these three seemingly disparate quantities obey formally analogous fluctuation theorems.

The obvious underlying link between the state transformations in these distinct contexts is the central role played by the majorisation criteria in the corresponding resource

theories. This leads to the primary open question of whether majorisation is necessary or just sufficient for the emergence of fluctuation theorems. The present investigation also suggests that there may be more operational contexts in quantum mechanics and beyond for which specific resources can be considered to fluctuate. In this respect, a fascinating problem is whether one might establish a hierarchy of fluctuating resources, whereby some fundamental quantity which obeys fluctuation theorems can be shown to induce a similar behaviour onto other resources. Quantum coherence is in fact an essential ingredient not only for entanglement, but also for more general non-classical correlations [216, 7]. It could be worthwhile to address whether fluctuation relations for the latter can be suitably derived by adapting and extending the present work. We further remark that all these phenomena can be naturally quantified by means of extensive quantities, based on the von Neumann and Rényi entropies, but generalisations of our results to non-extensive settings [2], e.g. adopting quantifiers based on Tsallis entropies, could also potentially be considered.

It is intriguing that an ancillary system to transfer and store coherence is necessary in the protocol we introduced, just like a conventional battery (a system that is able to store and transfer work, such as a weight or a piston) is used in thermodynamics. However, unlike work in classical thermodynamics, which can be determined by just measuring the difference in energy between the initial and final state, here to extract the distribution of fluctuating coherence e.g. according to the measurement protocol in Figure. 3.1(b), one irremediably destroys the initial superposition. It is this very uncertainty associated to fluctuations of coherence in the battery that ultimately allows the implementation of state transformations which would otherwise be forbidden to occur.

Further comparisons can be made between the fields. Because of the definition of the majorisation criteria for the state transformation (3.1), which map the diagonal elements of the initial state to those of the final state (rather than vice versa), the probability distribution in Eq. (3.29) is formed backwards, which means that the direction of the protocol giving rise to such a distribution is in contrast to the thermodynamic scenario. While the mathematical origin of this discrepancy is clear, its deeper physical meaning remains elusive. As remarked earlier in the text, from the analogue fluctuation theorems in Eqs. (3.40) and (3.44), it can also be seen that the partition function of statistical mechanics is akin to the diagonal rank of states in coherence theory, as both quantifiers affect the availability of the given resource, i.e., work in thermodynamics and coherence in this chapter.

The results in our study also have important implications for the resource theory of quantum coherence in its own right [24, 216]. In this respect, it is worth comparing explicitly the power of different assisted and unassisted scenarios for coherence distillation

and state manipulations under incoherent operations.

In the standard unassisted scenario, as depicted in Figure 3.1(a), the majorisation criteria recalled in Section 3.2 can be equivalently formulated in terms of relative entropy: the state transformation $|\psi\rangle_A \to |\phi\rangle_A$ is possible by deterministic IO (or SIO) if and only if [76, 77, 242, 55, 54, 56, 254]

$$C_{\mathrm{rel}}(\psi_A) \geq C_{\mathrm{rel}}(\phi_A). \tag{3.45}$$

If we consider instead the framework where the d-dimensional system A is assisted by an exact catalyst B, whose state τ_B needs to be returned unchanged, then the state transformation $|\psi\rangle_A \otimes |\tau\rangle_B \to |\phi\rangle_A \otimes |\tau\rangle_B$ is possible by deterministic IO if and only if [49]

$$\frac{S_\alpha\big(\Delta(\psi_A)\big) - \ln d}{|\alpha|} > \frac{S_\alpha\big(\Delta(\phi_A)\big) - \ln d}{|\alpha|}, \quad \forall\, \alpha \in (-\infty, \infty), \tag{3.46}$$

where the quantities $S_\alpha(\rho) := \mathrm{sgn}(\alpha) \ln\big(\mathrm{Tr}(\rho^\alpha)\big)/(1 - \alpha)$ denote a family of Rényi entropies.

In the paradigm investigated in this chapter, illustrated in Figure 3.1(b), the system A and the battery B are still prepared in an initial product state $|\psi\rangle_A \otimes |\lambda\rangle_B$, but the coherence in the battery B is allowed to change by an amount w with probability $P(w)$, so that the battery plays the role of an approximate catalyst. In this case, we have shown that *any* transformation of the form (3.1) can be implemented by IO or SIO on the system and the battery, provided the extracted coherence obeys the second law (3.34),

$$C_{\mathrm{rel}}(\psi_A) - C_{\mathrm{rel}}(\phi_A) \geq \langle w \rangle. \tag{3.47}$$

This is somehow comparable to the quantum thermodynamic setting recently investigated in [171], in which by allowing the buildup of arbitrarily small correlations between a system and a catalyst during a thermal operation, one finds that state transformations (for states diagonal in the energy eigenbasis) are specified by the second law expressed just in terms of the conventional Helmholtz free energy [171], rather than in terms of a whole family of Rényi entropies [44].

In the standard resource theory of coherence, it is currently understood that pure state transformations are reversible only in the asymptotic setting of many copies [216]. However, if a coherence battery is employed, then according to the coherence fluctuation theorem in Eq. (3.33) a single copy transformation becomes reversible, when accompanied by a nontrivial fluctuation in coherence with probability $P(w)$. This means that, if there exists a coherence fluctuation $w = \ln q_i - \ln p_i$, then there exists a reverse process with

equal and opposite coherence $w_{\mathrm{rev}} = \ln p_i - \ln q_i$, where the forward protocol is nevertheless exponentially preferred according to the analogue Crooks relation in Eq. (3.44).

Within the standard (unassisted) coherence resource theory, it is also known that in the asymptotic setting the distillable coherence under IO for any state is given by the relative entropy of coherence [242]. However, comparing Eqs. (3.45) and (3.47), we may conclude that the second law of coherence demonstrates that, although on average the distillable coherence of a state is given by the relative entropy, in the battery assisted framework there is an exponentially suppressed regime in which more coherence can be extracted.

As a next step, it would be desirable to generalise our analysis to the non-asymptotic regime and explore the role of coherence fluctuation relations in the context of one-shot state transformations under different classes of incoherent operations, following recent work on one-shot coherence dilution and distillation [253, 197], and inspired by thermodynamic studies of one-shot dissipated work [238, 109].

We have not considered the alternative assisted framework where the initial state ρ_{AB} contains correlations between system A and ancilla B, as that case would require the state of the system to be mixed. It is known that, in such a collaborative context, the asymptotic distillable coherence on A under local incoherent-quantum operations amounts to $S(\Delta(\rho_A))$, which yields a net gain over the unassisted case $C_{\mathrm{rel}}(\rho_A)$ by a quantity equal to the reduced von Neumann entropy $S(\rho_A)$ [58]. It may be worthy in the future to investigate coherence fluctuations in the battery, enhanced by initial correlations with the system. This would be especially interesting in view of the fact that the laws of thermodynamics in the presence of correlations (which might allow for seemingly paradoxical feats such as anomalous heat flow) have only recently begun to be understood in terms of physical processes [35].

It is hoped that, in the same way that the fluctuation theorems of statistical mechanics and thermodynamics have opened up a wide range of theoretic and experimental investigations, the fluctuation theorems of quantum coherence may inspire the discovery of new phenomena within coherence theory and applications, and beyond. It has already been shown in this chapter that fluctuating coherence allows one to break current limitations on reversibility and distillation in state transformations. The study of hybrid frameworks whereas a coherence battery may be employed to assist state transformations in different resource theories, such as athermality, entanglement, and more general manifestations of non-classicality, also deserves further investigation. This could complement recent studies of catalytic coherence for work extraction [1, 144] and reveal new crossing points between the characterisation of coherence and quantum correlations [218, 254] in quantum information theory.

Chapter 4

Indistinguishability and thermodynamics: a quantum Gibbs' paradox

4.1 Introduction

As explored in chapter 1, thermodynamics has always been inextricably linked with the abstract concept of information. Such connections have proven essential for solving paradoxes in a variety of thought experiments, notably including Maxwell's demon [30] and Loschmidt's paradox [120]. This integration between classical thermodynamics and information is also one of the main motivating factors in extending the theory to the quantum realm, where information held by the observer plays a similarly fundamental role [39].

In this chapter, we study the transition from classical to quantum thermodynamics in the context of the Gibbs paradox [97, 154, 10]. This thought experiment considers two gases on either side of a box, separated by a partition and with equal volume and pressure on each side. If the gases are identical, then the box is already in thermal equilibrium, and nothing changes after removal of the partition. If the gases are distinct, then they mix and expand to fill the volume independently, approaching thermal equilibrium with a corresponding entropy increase. The (supposed) paradox can be summarised as follows: what if the gases differ in some unobservable or negligible way – should we ascribe an entropy increase to the mixing process or not? This question sits uncomfortably with the view that thermodynamical entropy is an objective physical quantity.

Various resolutions have been described, from phenomenological thermodynamics to

statistical mechanics perspectives, and continue to be analysed [233, 68, 10]. A crucial insight by Jaynes [134] assuages our discomfort at the observer-dependent nature of the entropy change. For an *informed observer*, who sees the difference between the gases, the entropy increase has physical significance in terms of the work extractable through the mixing process – in principle, they can build a device that couples to the two gases separately (for example, through a semi-permeable membrane) and thus let each gas do work on an external weight independently. An *ignorant observer*, who has no access to the distinguishing degree of freedom, has no device in their laboratory that can exploit the difference between the gases, and so cannot extract work. For Jaynes, there is no paradox as long as one considers the abilities of the experimenter – a viewpoint central to the present work.

We study the Gibbs mixing process for quantum gases of identical bosons or fermions. This is motivated by recognising that the laws of thermodynamics must be modified to account for quantum effects such as coherence [160], which can lead to enhanced performance of thermal machines [234, 229]. The thermodynamical implications of identical quantum particles have received renewed interest for applications such as Szilard engines [193, 29], thermodynamical cycles [172, 237] and energy transfer from boson bunching [121]. Moreover, the particular quantum properties of identical particles, including entanglement, can be valuable resources in quantum information processing tasks [138, 170, 46]

We consider a toy model of an ideal gas with non-interacting quantum particles, distinguishing the two gases by a spin-like degree of freedom. We describe the mixing processes that can be performed by both informed and ignorant observers, taking into account their different levels of control, from which we can calculate the corresponding entropy changes and thus work extractable by each observer. For the informed observer, we recover the same results as obtained by classical statistical mechanics arguments. However, for the ignorant observer, there is a marked divergence from the classical case. Counterintuitively, the ignorant observer can typically extract more work from distinguishable gases – even though they appear indistinguishable – than from truly identical gases. In the continuum and large particle number limit which classically recovers the ideal gas, this divergence is maximal: *the ignorant observer can extract as much work from apparently indistinguishable gases as the informed observer.*

Our analysis hinges on the symmetry properties of quantum states under permutations of particles, defined in section 1.3. For the ignorant observer, these properties lead to non-trivial restrictions on the possible work extraction processes. Viewed another way, the microstates of the system described by the ignorant observer are highly non-classical entangled states. This implies a fundamentally different way of counting microstates, and therefore computing entropies, from what is done classically or even in semi-classical treat-
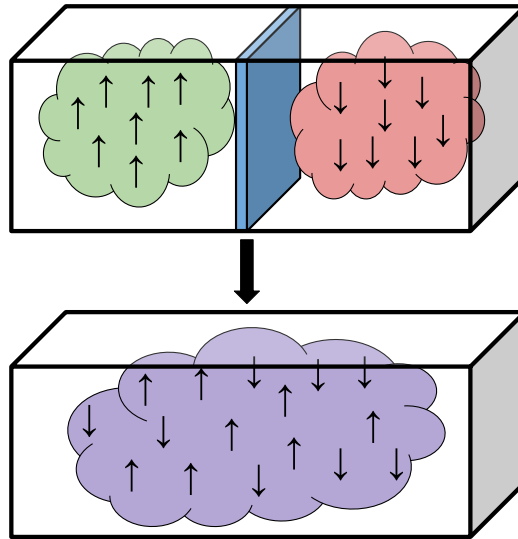
Figure 4.1: **The Gibbs paradox.** Two distinct gases of n particles at the same temperature and pressure are separated by a partition. This partition is removed and the gases are allowed to mix and reach equilibrium. Two observers calculating the entropy increase during the process disagree depending on their ability to distinguish the particles. An informed observer, who can measure the difference between the gases, calculates $2n \ln 2$, while an observer ignorant of the difference records no entropy change. In this chapter, we ask how the situation changes when classical particles are replaced by identical quantum particles.

ments of quantum gases. Therefore we uncover a genuinely quantum thermodynamical effect in the Gibbs mixing scenario.

4.2   Results

4.2.1   Set-up

We consider a gas of N particles inside a box, such that each particle has a position degree of freedom, denoted x^1, and a second degree of freedom which distinguishes the gases. Since we only consider the case of two types of gases, this is a two-dimensional degree of freedom and we refer to this as the "spin" s (although it need not be an intrinsic angular momentum). Classically, the two spin labels are $\uparrow, \downarrow$, and their quantum analogues are orthogonal states $|\uparrow\rangle, |\downarrow\rangle$.

---

[1]This 'position' degree of freedom is just used as a convenient d-dimensional degree of freedom. This could indeed be some discrete 'position' degree of freedom, in which case the normal conjugate methods of calculating other degrees of freedom such as momentum would apply, although such calculations are not necessary here.

Following the traditional presentation of the Gibbs paradox, the protocol starts with two independent gases on different sides of a box: n on the left and $m = N - n$ on the right (see Fig. 4.1). Each side is initially thermalised with an external heat bath B at temperature T.

In our toy model, each side of the box consists of $d/2$ "cells" (d is even) representing different states that can be occupied by each particle. These states are degenerate in energy, such that the Hamiltonian of the particles vanishes. This might seem like an unrealistic assumption; however, this model contains the purely combinatorial (or "state-counting") statistical effects, first analysed by Boltzmann [40], that are known to recover the entropy changes for a classical ideal gas [68, 203, 73] using the principle of equal a priori probabilities. One could instead think of this setting as approximating a non-zero Hamiltonian in the high-temperature limit (a zero Hamiltonian being the subject of equation (1.49)), such that each cell is equally likely to be occupied in a thermal state. Since the particle number is strictly fixed, we are working in the canonical ensemble (rather than the grand canonical ensemble).

Work extraction can be modelled in various ways in quantum thermodynamics. In the resource-theoretic approach based on thermal operations [127, 43], as we have seen, one keeps track of all resources by treating the system (here, the particles), heat bath and work reservoir (or battery) as interacting quantum systems. The work reservoir is an additional system with non-degenerate Hamiltonian, the non-degeneracy being important as we relate said systems energy changes with work done by or on the system (generalising the classical idea of a weight being lifted and lowered). The gases on either side of the box start in a state of local equilibrium and via mixing approach global equilibrium. We therefore consider the extractable work to be given by the difference in non-equilibrium free energy F [36] between initial and final states, where $F(\rho) = \langle E \rangle_\rho - k_B T S(\rho)$, $\langle E \rangle_\rho = \mathrm{tr}(\rho H)$ being the mean energy (zero in our case) and $S(\rho) = -\mathrm{tr}(\rho \ln \rho)$ the von Neumann entropy in natural units. The extractable work in a process that takes ρ to ρ' is then

$$W \leq F(\rho) - F(\rho') = k_B T \left[S(\rho') - S(\rho) \right]. \tag{4.1}$$

In a classical reversible process, the extractable work is equal to the change in free energies. This is generally an over-simplification for small systems, in which work can be defined in various ways [178] – e.g. required to be deterministic in the resource theory context [127] or as a fluctuating random variable [4, 65], requiring consideration of other varieties of free energy. However, equation (4.1) will turn out to be sufficient for our purposes in the sense of mean extractable work. We find the inequality to be saturable using thermal operations and characterise fluctuations around the mean in the latter part of our results section.

Table 1: Summary of the observers' abilities		
Observer	**Can**	**Can't**
Informed	Access the spin and spatial degrees of freedom	Change the number of up or down spins
Ignorant	Access the spatial degree of freedom	Access the spin degree of freedom

Our analysis compares the work extracted by two observers with different levels of knowledge: the *informed observer*, who can tell the difference between the two gases, and the *ignorant observer*, who cannot. The difference between these observers is that the former has access to the spin degree of freedom s, whereas the latter does not (summarised in Table 1).

It is important to point out that, for the informed observer, the spin acts as a "passive" degree of freedom, meaning that it can be measured but not actively changed. In other words, the two types of gases cannot be converted into each other. This assumption is always implicitly present in discussions of the Gibbs paradox – without it, the distinguishing degree of freedom would constitute another subsystem with its own entropy changes. One could also describe the spin as an *information-bearing degree of freedom* [151]. The question is whether the information encoded within the spin state has an impact upon the thermodynamics of mixing.

4.2.2   Classical case

Classically, the microstates described by the informed observer are specified by counting how many particles exist with each position x and spin s – since the particles are indistinguishable [16]. The ignorant observer has a different state space given by coarse-graining these states – the classical equivalent of "tracing out" the spin degree of freedom. Thus the ignorant observer can extract only as much work from two different gases as from a single gas, recovering Jaynes' original statement [134]. These intuitively obvious facts are shown by a formal construction of the state spaces in appendix C.1. Paralleling our later quantum treatment, this establishes that the classical and quantum cases can be compared fairly.

The amount of extractable work in the classical case can be straightforwardly argued by state counting. Consider the gas initially on the left side – the number of ways of

distributing n particles among $d/2$ cells is $\binom{n+d/2-1}{n}$. In the thermal state, each configuration occurs with equal probability. Therefore the initial entropy, also including the gas on the right, is $\ln\binom{n+d/2-1}{n} + \ln\binom{m+d/2-1}{m}$. For distinguishable gases, each gas can deliver work independently, with an equal distribution over $\binom{n+d-1}{n}\binom{m+d-1}{m}$ configurations. For indistinguishable gases, the final thermal state is described as an equal distribution over all ways of putting $N = n + m$ particles into d cells, of which there are $\binom{N+d-1}{N}$. Hence the entropy change in each case is

$$
\begin{aligned}
\Delta S = {} & \ln\binom{n+d-1}{n} + \ln\binom{m+d-1}{m} - \ln\binom{n+d/2-1}{n} \\
& - \ln\binom{m+d/2-1}{m} \quad \text{(distinguishable)},
\end{aligned} \tag{4.2}
$$

$$
\begin{aligned}
\Delta S = {} & \ln\binom{N+d-1}{N} - \ln\binom{n+d/2-1}{n} \\
& - \ln\binom{m+d/2-1}{m} \quad \text{(indistinguishable)}.
\end{aligned} \tag{4.3}
$$

Note that $\Delta S \neq 0$ even in the indistinguishable case, which may seem at odds intuitively with the result for an ideal gas. However, one can check that $\Delta S = \mathcal{O}(\ln N)$ in the limit of large d (whereby the box becomes a continuum) and large N. This is negligible compared with the ideal gas expression of $N \ln 2$ for distinguishable gases [88] [2]. (Due to a subtle technicality with classical identical particles, formulas (4.2),(4.3) might be regarded as upper bounds to the true values – see appendix C.1.) Note that a classical analogue of fermions can be made by importing the Pauli exclusion principle, so that two or more particles can never occupy the same cell. This has the effect of replacing the binomial coefficients of the form $\binom{N+d-1}{N}$ in (4.2) and (4.3) by $\binom{d}{N}$.

4.2.3 Quantum case

Informed observer

Compared with the classical case, we must be more explicit about the role of the spin s as a "passive" degree of freedom for the informed observer. This observer may obtain information about the numbers of spin-↑ and spin-↓ particles. Thus they can engineer spin-dependent operations conditional on these numbers, but cannot change the number of each spin.

For identical gases, the result is of course the same as for the ignorant observer, and

---

[2]See [68, p. 43] for a more detailed discussion of this approximation.
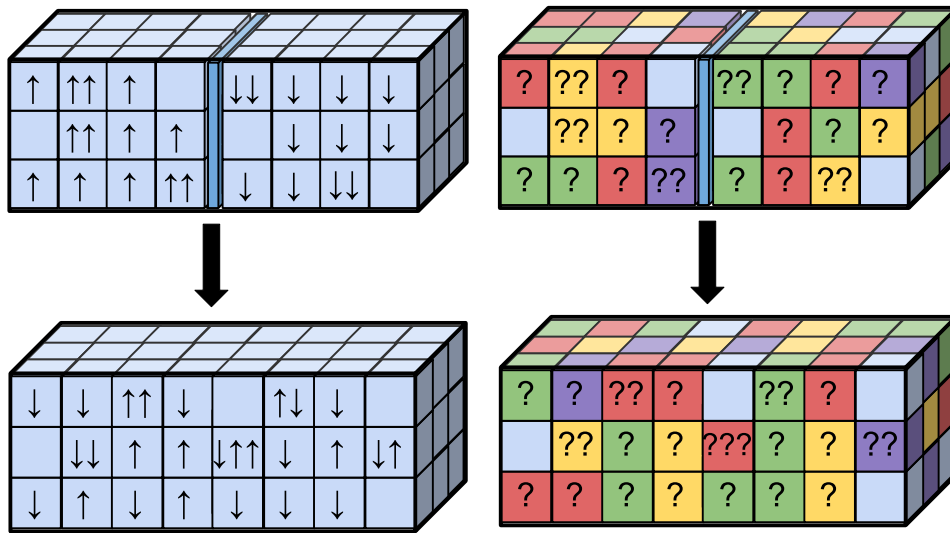
Figure 4.2: **Schematic of the quantum mixing process.** Two diagrams representing the mixing of indistinguishable (bosonic) quantum gases from the perspective of the *informed* (left) and *ignorant* (right) observers. Initially, n spin-↑ particles are found on the left and m spin-↓ on the right. The particles are then allowed to mix while coupling to an external heat bath and work reservoir. The informed observer describes microstates via the number of particles in each cell, and their respective spins. The ignorant observer cannot tell the spins states, but describes microstates (schematically depicted here by different colours) as superpositions of cell configurations, determined by the decomposition (4.6).

the classical case (4.3). For distinguishable gases, each gas behaves as an independent subsystem; thus, the entropy changes are the same as for classical distinguishable gases (4.2).

The remainder of this section is devoted to the ignorant observer, for which we find a departure from the classical case.

Hilbert space

The peculiarities of the quantum case stem from a careful look at the Hilbert space structure. The Hilbert space of a single particle is a product $\mathcal{H}_1 = \mathcal{H}_x \otimes \mathcal{H}_s$ of a part for the spatial degree of freedom x and a part for the spin s. Since there are d cell modes and two spin states, these parts have dimensions $\dim \mathcal{H}_x = d$, $\dim \mathcal{H}_s = 2$. For N distinguishable particles, the state space would be $\mathcal{H}_1^{\otimes N}$. However as introduced in section 1.3, for bosons and fermions, which are quantum indistinguishable particles, states lie in the symmetric and antisymmetric subspaces, respectively (in first quantisation). This symmetry refers to the wavefunction under permutations of particles: for bosons, there is no change, whereas for fermions, each swap of a pair incurs a minus sign in the global phase. The physical

Hilbert space of N particles can then be written as

$$\mathcal{H}_N = P_\pm \left(\mathcal{H}_x^{\otimes N} \otimes \mathcal{H}_s^{\otimes N} \right), \tag{4.4}$$

where $P_{+(-)}$, defined in equation 1.26, is the projector onto the (anti-)symmetric subspace.

Since each particle carries a position and spin state, a permutation Π of particles is applied simultaneously to these two parts: Π acts on the above Hilbert space in the form $\Pi_x \otimes \Pi_s$. The requirement of an overall (anti-)symmetric wavefunction effectively couples these two degrees of freedom via their symmetries. For a familiar example, consider two particles. The spin state space can be broken down into the symmetric "triplet" subspace spanned by $|\uparrow\uparrow\rangle$, $|\downarrow\downarrow\rangle$ and $|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle$, and the antisymmetric "singlet" subspace consisting of $|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle$. For bosons, overall symmetry requires that a triplet spin state be paired with a symmetric spatial wavefunction, and a singlet spin state with an antisymmetric spatial function. For fermions, opposite symmetries are paired.

With more particles, the description is more complex, but the main idea of paired symmetries remains the same. Following [5], our main tool is Schur-Weyl duality [102], which decomposes

$$\mathcal{H}_x^{\otimes N} = \bigoplus_\lambda \mathcal{H}_x^\lambda \otimes \mathcal{K}_x^\lambda, \tag{4.5}$$

where λ runs over all Young diagrams of N boxes and no more than d rows [3]. In technical terms, $\mathcal{H}_x^\lambda$ and $\mathcal{K}_x^\lambda$ carry irreducible representations of the unitary group $\mathrm{U}(d)$ and the permutation group S_N of N particles, respectively. More concretely, a non-interacting unitary operation on the positions of all the particles, $u_x^{\otimes N}$, is represented in the decomposition (4.5) as an independent rotation within each of the $\mathcal{H}_x^\lambda$ spaces. The term "irreducible" refers to the fact that each of these spaces may be fully explored by varying the unitary u_x. Similarly, a permutation of the particles in the spatial part of the wavefunction is represented by an action on each $\mathcal{K}_x^\lambda$ space. Thus each block labelled by λ in the decomposition (4.5) has a specific type of permutation symmetry.

The same decomposition works for the spin part $\mathcal{H}_s^{\otimes N}$. However, since this degree of freedom is two-dimensional, each λ is constrained to have no more than two rows. We can think of s as describing a total angular momentum formed of N spin-1/2 particles, and in fact λ can be replaced by a total angular momentum eigenvalue J varying over the range $N/2, N/2 - 1, \ldots$.

After putting the spatial and spin decompositions together, projecting onto the overall (anti-)symmetric subspace causes the symmetries of the two parts to be linked. For

---

[3]A Young diagram can be described simply by a non-increasing set of ($\leq d$) positive integers summing up to N.

bosons, the λ label for x and s must be the same; for fermions, they are transposes of each other (i.e, related by interchanging rows and columns). This results in the form

$$\mathcal{H}_N = \bigoplus_\lambda \mathcal{H}_x^\lambda \otimes \mathcal{H}_s^\lambda \quad \text{for bosons,}$$

$$\mathcal{H}_N = \bigoplus_\lambda \mathcal{H}_x^{\lambda^T} \otimes \mathcal{H}_s^\lambda \quad \text{for fermions.} \tag{4.6}$$

Instead of the label λ, from now on we use the angular momentum number J and generally write this decomposition as $\bigoplus_J \mathcal{H}_x^J \otimes \mathcal{H}_s^J$ – bearing in mind that $\mathcal{H}_x^J$ is different for bosons and fermions. In terms of the earlier $N = 2$ example, $J = 1$ corresponds to the spin triplet subspace, and $J = 0$ to the spin singlet.

Another way of describing the decomposition (4.6) is that it provides a convenient basis $|J, q\rangle_x |J, M\rangle_s |\phi_J\rangle_{xs}$, known as the Schur basis [112]. Here, $\{|J, q\rangle_x\}_q$ is a basis for $\mathcal{H}_x^J$ and $\{|J, M\rangle_s\}_M$ a basis for $\mathcal{H}_s^J$. $M = -J, -J+1, \ldots, J$ can be interpreted as the total angular momentum quantum number along the z-axis. $|\phi_J\rangle_{xs} \in \mathcal{K}_x^J \otimes \mathcal{K}_s^J$ is a state shared between the x and s degrees of freedom.

Thermalisation for ignorant observer

Since the ignorant observer cannot interact with spin, their effective state space is described by tracing out the factor $\mathcal{H}_s$ for each particle. In terms of the decomposition (4.6) and corresponding basis described above, this means that an initial density matrix ρ, after tracing out s, is of the form

$$\rho_x := \mathrm{tr}_s \rho = \bigoplus_J p_J \rho_x^J \otimes \mathrm{tr}_s |\phi_J\rangle\langle\phi_J|_{xs}, \tag{4.7}$$

where ρ_x^J is a density matrix on $\mathcal{H}_x^J$, occurring with probability p_J. Note that there is no coherence between different values of J, and that the components ρ_x^J are mutually perfectly distinguishable by a measurement of their J.

Additionally, the allowed operations must preserve the bosonic or fermionic exchange symmetry. Any global unitary U_{xBW}, coupling the spatial degree of freedom of the particles to the heat bath and work reservoir, must therefore commute with permutations on the spatial part: $[U_{xBW}, \Pi_x] = 0$ for all Π. By Schur's Lemma, such a unitary decomposes as $U = \bigoplus_J U^J \otimes I^J$, where U^J operates on the $\mathcal{H}_x^J$ component, with an identity I^J on $\mathcal{K}_x^J$. Hence each J component is operated upon independently, the spin eigenvalue J being conserved.

In summary, therefore, the ignorant observer may engineer any thermal operation extracting work separately from each J component (depicted in Fig. 4.2). We can think of their operations being conditioned on the spatial symmetry type, and although J is observed to fluctuate randomly, a certain amount of work is extracted for each J (see the latter part of the results section for a more detailed analysis of this fluctuation) [4]. For each J, there exists an operation within the thermal operations framework [127] that performs deterministic work extraction saturating inequality (4.1). This is because the transformation is between (energy-degenerate) uniformly mixed states of differing dimension.

The question of optimal work extraction thus reduces to calculating the entropy of the initial state (4.7) and finding the maximum entropy final state. The fully thermalised final state seen by the ignorant observer is maximally mixed within each J block:

$$\rho'_x = \bigoplus_J p_J \frac{I^J_x}{d_J} \otimes \mathrm{tr}_s |\phi_J\rangle\langle\phi_J|_{xs}, \tag{4.8}$$

where I^J_x is the identity on $\mathcal{H}^J_x$ and d_J is the corresponding dimension.

The overall entropy change is the average over all J, found to be (with details in appendix C.2):

$$\begin{aligned} \Delta S_{\mathrm{igno}} &= \sum_J p_J \Delta S^J_{\mathrm{igno}}, \\ &= \sum_J p_J \ln d^B_J - \ln \binom{n + d/2 - 1}{n} - \ln \binom{m + d/2 - 1}{m} \end{aligned} \tag{4.9}$$

for bosons, and

$$\Delta S_{\mathrm{igno}} = \sum_J p_J \ln d^F_J - \ln \binom{d/2}{n} - \ln \binom{d/2}{m} \tag{4.10}$$

for fermions. Expressions for the dimensions $d^{B,F}_J$ are found in appendix C.4:

$$\begin{aligned} d^B_J &= \frac{(2J+1)\left(\frac{N}{2} - J + d - 2\right)!\left(\frac{N}{2} + J + d - 1\right)!}{\left(\frac{N}{2} - J\right)!\left(\frac{N}{2} + J + 1\right)!(d-1)!(d-2)!}, \\ d^F_J &= \frac{(2J+1)d!(d+1)!}{\left(\frac{N}{2} + J + 1\right)!\left(\frac{N}{2} - J\right)!\left(d - \frac{N}{2} + J + 1\right)!\left(d - \frac{N}{2} - J\right)!}. \end{aligned} \tag{4.11}$$

The probabilities p_J are found (see appendix C.2) from the Clebsch-Gordan coefficients

---

[4]Note that the work extraction process does not involve a measurement by the observer – only a coupling to the apparatus that depends on the value of J. Therefore there is no need to consider an additional entropic measurement cost, unlike the case of Maxwell's demon [255, 30]

Table 2: Summary of results

	Quantum	Classical	Quantum	Quantum	Classical
Limit	(no limit)	(no limit)	$(d \gg n^2)$	$(d \gg n^2 \gg 1)$	$(d \gg n^2 \gg 1)$
ΔS_{info}	$2\ln\binom{n+d-1}{n} - 2\ln\binom{n+d/2-1}{n}$	$2\ln\binom{n+d-1}{n} - 2\ln\binom{n+d/2-1}{n}$	$\cdots$	$\approx 2n\ln 2$	$\approx 2n\ln 2$
ΔS_{igno}	$\sum_J p_J \ln d_J^B - 2\ln\binom{n+d/2-1}{n}$	$\ln\binom{2n+d-1}{2n} - 2\ln\binom{n+d/2-1}{n}$	$\approx \Delta S_{\text{info}} - H(\boldsymbol{p}) - \frac{n^2}{2d^2}$	$\approx 2n\ln 2$	≈ 0

Entropy changes $\Delta S_{\text{info}}, \Delta S_{\text{igno}}$ for the informed and ignorant observers and their limits are expressed for bosons with $n = m$. For fermions, replace the dimension of the symmetric subspace $\binom{n+d-1}{n}$ with that of the antisymmetric one $\binom{d}{n}$ and d_J^B by d_J^F (both of which are defined in equation (4.11)).

$C(j_1, m_1; j_2, m_2; J, M)$ describing the coupling of two spins with angular momentum quantum numbers (j_1, m_1), (j_2, m_2) into overall quantum numbers (J, M). Here, the two spins are the groups of particles on the left and right, respectively.

For identical gases, all particles have spins in the same direction, so the spin wavefunction is simply $|\uparrow\rangle^{\otimes N}$. This state lies fully in the subspace of maximal total spin eigenvalue, $J = M = N/2$ – which is also fully symmetric with respect to permutations. Thus the spin part factorises out (i.e., there is no correlation between spin and spatial degrees of freedom). It is then clear that dimension counting reduces to the classical logic of counting ways to distribute particles between cells. Indeed, the dimension of the subspace $\mathcal{H}_x^{N/2}$ is $d_{N/2}^B = \binom{N+d-1}{N}$ for bosons and $d_{N/2}^F = \binom{d}{N}$ for fermions. It follows that we recover the entropy as the classical case of indistinguishable particles (4.3).

For orthogonal spins, there are n spin-$\uparrow$ and m spin-$\downarrow$, leading to $M = (n - m)/2$ and a distribution over different values of J according to

$$p_J = \frac{(2J + 1)n!m!}{\left(\frac{N}{2} + J + 1\right)! \left(\frac{N}{2} - J\right)!}. \tag{4.12}$$

The resulting entropies and significant limits are discussed after an example.

Example. Taking $n = m = 1$ demonstrates the mechanism behind the state space decomposition. For two particles, there are only two values of J, corresponding to the

familiar singlet and triplet subspaces:

$$
\begin{aligned}
\mathcal{H}_s^0 &= \mathrm{span}\left\{ |\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle \right\}, \\
\mathcal{H}_s^1 &= \mathrm{span}\left\{ |\uparrow\uparrow\rangle, |\downarrow\downarrow\rangle, |\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle \right\}.
\end{aligned}
\tag{4.13}
$$

Consider a spatial configuration where a spin-$\uparrow$ particle is on the left in cell i, and a spin-$\downarrow$ is on the right in cell j. For bosons, the properly symmetrised wavefunction is

$$
\begin{aligned}
|\psi_{i,j}\rangle &:= \frac{1}{\sqrt{2}} \left(|i_L j_R\rangle_x |\uparrow\downarrow\rangle_s + |j_R i_L\rangle_x |\downarrow\uparrow\rangle_s \right) \\
&= \frac{1}{\sqrt{2}} \left[\frac{|i_L j_R\rangle - |j_R i_L\rangle}{\sqrt{2}} \cdot \frac{|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle}{\sqrt{2}} \quad (J=0) \right. \\
&\qquad \left. + \frac{|i_L j_R\rangle + |j_R i_L\rangle}{\sqrt{2}} \cdot \frac{|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle}{\sqrt{2}} \quad (J=1) \right].
\end{aligned}
\tag{4.14}
$$

So $p_0 = p_1 = 1/2$, and the spatial component of this state is conditionally pure for both J. The initial thermal state is a uniform mixture of all such $|\psi_{i,j}\rangle$, with $(d/2)^2$ terms. Thus $S(\rho_x^0) = S(\rho_x^1) = 2(\ln d - \ln 2)$. For the final thermal state, we observe that

$$
\begin{aligned}
\mathcal{H}_x^0 &= \mathrm{span}\left\{ |ij\rangle - |ji\rangle \mid i < j \right\}, \\
\mathcal{H}_x^1 &= \mathrm{span}\left\{ |ij\rangle + |ji\rangle \mid i \le j \right\},
\end{aligned}
\tag{4.15}
$$

where i, j now label cells either on the left or right. The corresponding dimensions are $d_0 = d(d-1)/2$, $d_1 = d(d+1)/2$. Within the $J = 0$ subspace, the entropy change is $\ln[d(d-1)/2] - 2\ln d + 2\ln 2 = \ln(1 - 1/d) + \ln 2$, and for $J = 1$, it is $\ln[d(d+1)] - 2\ln d + 2\ln 2 = \ln(1 + 1/d) + \ln 2$. Overall, therefore,

$$
\begin{aligned}
\Delta S_{\mathrm{igno}} &= \frac{1}{2} \ln\left(1 - \frac{1}{d} \right) + \frac{1}{2} \ln\left(1 + \frac{1}{d} \right) + \ln 2 \\
&= \frac{1}{2} \ln\left(1 - \frac{1}{d^2} \right) + \ln 2.
\end{aligned}
\tag{4.16}
$$

For the informed observer, we have $\Delta S_{\mathrm{info}} = 2\ln 2$. For identical gases, we find $\Delta S_{\mathrm{iden}} = \ln(1 + 1/d) + \ln 2$, strictly greater than ΔS_{igno}, but the two become equal in the limit $d \to \infty$.

Repeating the same calculation with fermions, the symmetric and antisymmetric states now pair up oppositely. Then ΔS_{igno} is the same as for bosons. However, we have $\Delta S_{\mathrm{iden}} = \ln(1 - 1/d) + \ln 2 < \Delta S_{\mathrm{igno}}$. Unlike for bosons, two distinguishable fermions permit more extractable work by the ignorant observer than two identical fermions.
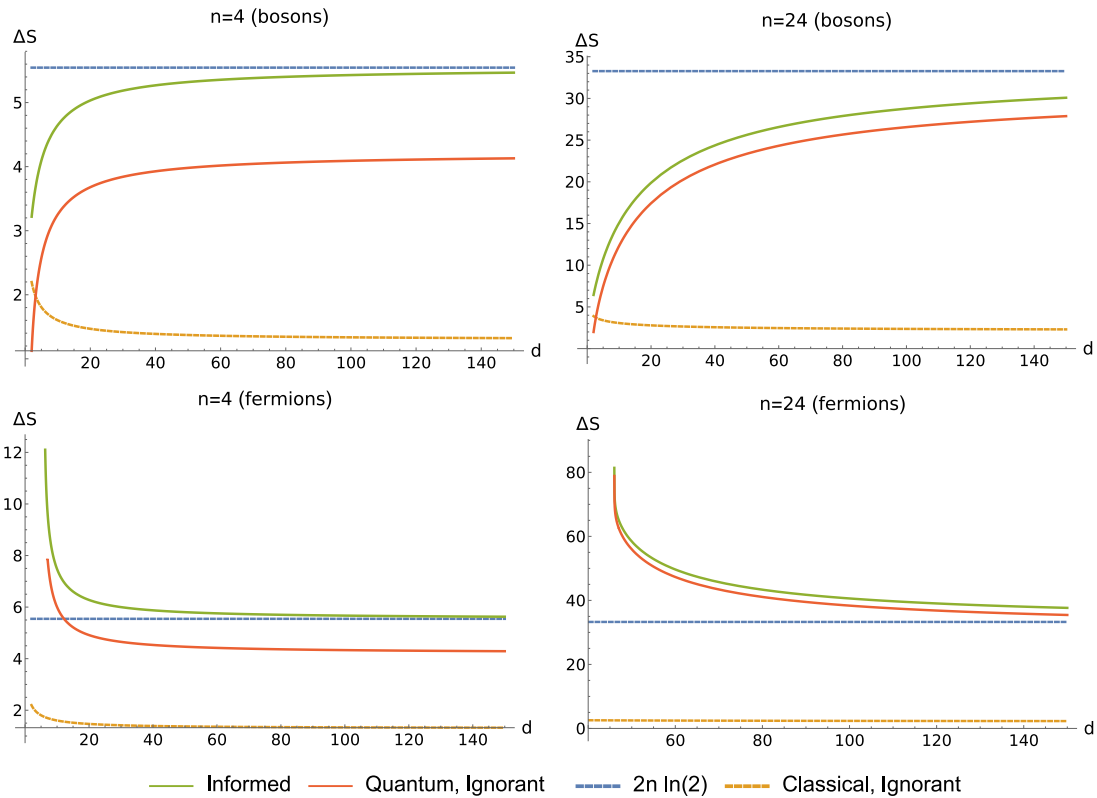
Figure 4.3: **Entropy changes as a function of dimension.** Series of plots showing $\Delta S_{\text{info}}, \Delta S_{\text{igno}}$ against the total cell number d of the system. Figures in the top row are for bosonic systems of differing particle number n and figures in the bottom row show the same for fermionic systems. Note that we have taken the initial number of particles on either side of the box to be equal, $n = m$ in all cases. For comparison, all four figures also display the classical changes in entropy for an informed/ignorant observer. The behaviour of the deficit between ΔS for an informed/ignorant observer of quantum particles agrees with the *low density limit* in equation (4.17) where we can see ΔS_{info} tending to the classical limit $2n \ln(2)$ with ΔS_{igno} trailing behind by a deficit of $n^2/d^2 + H(\boldsymbol{p})$. Additionally, by comparing the different plots, we can see the low-dimensional fermionic advantage where the change in entropy is even greater than the classical $2n \ln(2)$ value.

4.2.4   Entropy changes and limits

In Fig. 4.3 we plot both ΔS_{info} and ΔS_{igno} as a function of dimension for bosons and fermions. Below we analyse the special cases and limits which emerge from these expressions, summarised in Table 2.

Special cases

With bosons, there are two special cases in which it is easily proven that distinguishable gases are less useful than indistinguishable ones for the ignorant observer. The first case is the example above, with $n = m = 1$. In addition, for $d = 2$, we have $d_J^B = 2J + 1$ – so

the largest subspace is that with maximal $J = N/2$. The largest entropy change is then obtained when $p_{N/2} = 1$, which is satisfied precisely for indistinguishable gases.

For fermions, we see from Fig. 4.3 that the greatest work – for both observers – is obtained for small d. An intuitive explanation is that the Pauli exclusion principle causes the initial state to be constrained and thus have low entropy. For example, with the minimal dimension $d = 2n = 2m$, we have $\Delta S_{\text{info}} = 2 \ln \binom{2n}{n} \approx 4n \ln 2$ to leading order when n is large. The ignorant observer can do almost as well: the state is entirely contained in the $J = 0$ subspace, with $d_0^F = \frac{(2n)!(2n+1)!}{(n!)^2(n+1)!^2} = \frac{2n+1}{(n+1)^2}\binom{2n}{n}^2$, giving $\Delta S_{\text{igno}} \approx 4n \ln 2$ for large n. This is twice as much as for the classical ideal gas.

Low density limit

The most interesting conclusion is reached in the limit of large $d \gg n^2$, which we term the *low density limit*. For simplicity, we take $n = m$. To lowest order in n^2/d, we find

$$\Delta S_{\text{igno}} \approx \Delta S_{\text{info}} - H(\boldsymbol{p}) - \frac{n^2}{2d^2}, \tag{4.17}$$

where $H(\boldsymbol{p}) = -\sum_J p_J \ln p_J$ is the Shannon entropy of the distribution p_J. Thus, as $d \to \infty$, the ignorant observer can extract as much work as the informed one, minus an amount $H(\boldsymbol{p})$. This gap is evident from the graphs in Fig. 4.3.

Now consider the limit $d \gg n^2$, $n \gg 1$, with both low density and large particle number. Classically, this limit recovers ideal gas behaviour – the large dimension limit can be thought of as letting the box become a continuum. In appendix C.6, we show that $H(\boldsymbol{p})$ (which depends only on n, not d), behaves as

$$H(\boldsymbol{p}) \approx \frac{1}{2} \ln n + 0.595..., \tag{4.18}$$

with a correction going to zero as $n \to \infty$. Recall that the entropy change for the informed observer is approximately $2n \ln 2$ in this limit. Therefore the deficit $H(\boldsymbol{p})$, which is logarithmic, becomes negligible compared with $2n \ln 2$. Thus the ignorant observer can extract essentially as much work as the informed observer: $\Delta S_{\text{igno}} \approx \Delta S_{\text{info}} \approx 2n \ln 2$. This result is remarkable because it shows an extreme departure from the classical case in the macroscopic limit.

Explaining the low density limit

An important feature of the low density limit is that the final entropy becomes as large as it could possibly be: ρ'_x becomes maximally mixed over its whole state space. This is true for any N, not just large numbers. We now give an explanation of this phenomenon, which proceeds by counting the number of mutually orthogonal states which can be accessed by the ignorant observer.

The important point about the low density limit is that particles almost never sit on top of each other – that is, almost all states are such that precisely N cells are occupied, each with a single particle. More formally, the number of ways of putting N bosonic particles into d cells is $\binom{N+d-1}{N} \approx \binom{d}{N}$ when d is large, where the approximation means the ratio of the two sides is close to unity. Let us refer to each of these $\binom{d}{N}$ choices of (singly) occupied cells as a *cell configuration*. For each cell configuration, there are $\binom{N}{n}$ *spin configurations*, i.e., ways of distributing the n spin-↑ and m spin-↓ particles. In classical physics, the ignorant observer cannot distinguish any of the spin configurations corresponding to a single cell configuration. In quantum mechanics, remarkably, there are precisely $\binom{N}{n}$ states which can be fully distinguished by the ignorant observer, each being a superposition of different spin configurations.

Let us choose a single cell configuration – without loss of generality, let cells $1, \ldots, N$ be occupied. The state of a spin configuration is denoted as a permutation of

$$|\uparrow\rangle_1 \cdots |\uparrow\rangle_n |\downarrow\rangle_{n+1} \cdots |\downarrow\rangle_N \in (\mathbb{C}^2)^{\otimes N}, \tag{4.19}$$

where each cell is treated as a qubit with basis states $|\uparrow\rangle, |\downarrow\rangle$ according to which type of spin occupies it. (Note that the subsystems being labelled are here are the occupied cells, not particles.)

Again using Schur-Weyl duality, the state space of N qubits can be decomposed as

$$(\mathbb{C}^2)^{\otimes N} = \bigoplus_J \mathcal{H}^J \otimes \mathcal{K}^J. \tag{4.20}$$

Due to this decomposition, there is a natural basis $|J, M, p\rangle$, where SU(2) spin rotations $u_s^{\otimes N}$ act on the M label (denoting the eigenvalue of the total z-direction spin), and permutations Π of the N cells act on the p label.

How do we represent the effective state seen by the ignorant observer? In the representation used here, this corresponds to *twirling* over the spin states, i.e., performing a Haar measure average over all spin rotations $u_s^{\otimes N}$ [23]. In the basis $|J, M, p\rangle$, however, this is a

straightforward matter of tracing out the $\mathcal{H}^J$ subspaces, since only these are acted on by the twirling operation. Thus the ignorant observer has access to states labelled as $|J, p\rangle$.

How much information has been lost by tracing out $\mathcal{H}^J$? In fact, none – the label $M = (n - m)/2$ is fixed. Therefore the experimenter can perfectly distinguish all the basis states $|J, p\rangle$ – and there are just as many of these as there are spin configurations, namely $\binom{N}{n}$.

For example, take $n = m = 1$: the two spin configurations are $|\uparrow\downarrow\rangle, |\downarrow\uparrow\rangle$, and for some pair of occupied cells, the two distinguishable states are

$$|J = 1, M = 0, p = 0\rangle = \frac{1}{\sqrt{2}} (|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle),$$

$$|J = 0, M = 0, p = 0\rangle = \frac{1}{\sqrt{2}} (|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle). \qquad (4.21)$$

Since these are respectively in the triplet and singlet subspaces, they remain orthogonal even after twirling. They can be distinguished by mixing the cells at a balanced beam splitter: it is easy to show that the symmetric state ends up with a superposition of both particles in cell 1 and both in cell 2, while the antisymmetric state ends up with one particle on each side. Therefore, after this beam splitter, the two states can be distinguished by counting the total particle number in each cell.

A slightly more complex example is with $n = 2, m = 1$. Then the distinguishable basis states for three occupied cells are

$$\left| J = \frac{3}{2}, M = \frac{1}{2}, p = 0 \right\rangle = \frac{1}{\sqrt{3}} (|\uparrow\uparrow\downarrow\rangle + |\uparrow\downarrow\uparrow\rangle + |\downarrow\uparrow\uparrow\rangle),$$

$$\left| J = \frac{1}{2}, M = \frac{1}{2}, p = 0 \right\rangle = \frac{1}{\sqrt{2}} (|\downarrow\uparrow\uparrow\rangle + |\uparrow\downarrow\uparrow\rangle), \qquad (4.22)$$

$$\left| J = \frac{1}{2}, M = \frac{1}{2}, p = 1 \right\rangle = \sqrt{\frac{2}{3}} |\uparrow\uparrow\downarrow\rangle - \frac{1}{\sqrt{6}} (|\uparrow\downarrow\uparrow\rangle + |\downarrow\uparrow\uparrow\rangle).$$

Observe that the argument in this section does not depend in anyway on the exchange statistics of the particles, explaining why we see the same limit for bosons and fermions.

Quantumness of the protocol

The above discussion of the low density limit clarifies the fundamental reason why the quantum ignorant observer performs better than the classical one. The distinguishable states comprising the final thermalised state are superpositions of different spin configu-

rations. We might describe a classical observer within the quantum setting as one who is limited to operations diagonal in the basis of cell configurations – that is, they are only able to count the number of particles occupying each cell. For such an observer, these superposition states are indistinguishable.

A crucial question is then: how difficult is it to engineer the quantum protocol for the ignorant observer? We can imagine that the heat bath and work reservoir might naturally couple to the system in the cell occupation basis (if this is the basis that emerges in the classical case). The required coupling is in the Schur basis $|J, q\rangle_x$, which are generally highly entangled between cells. A sense of their complexity is given by the unitary that rotates the Schur basis to the computational basis, known as the Schur transform. Efficient algorithms to implement this transform have been found [18], with a quantum circuit whose size is polynomial in $N, d, \ln(1/\epsilon)$, allowing for error ϵ. This circuit is related to the quantum Fourier transform, an important subroutine in many quantum algorithms. Thus, while the Schur transform can be implemented efficiently, it appears that engineering the required work extraction protocol – in the absence of fortuitous symmetries in the physical systems being used – may be as complex as universal quantum computation.

Work fluctuations

The work extraction protocol we have presented is not deterministic: for each value of J, a different amount of work is extracted with probability p_J. This is typically expected of thermodynamics of small systems; however, in classical macroscopic thermodynamics, such fluctuations are negligible. We can ask whether the same is true of the work extracted by the ignorant observer in the quantum case, especially in the low density and large particle number limits.

One informative way of quantifying the fluctuations is via the variance of entropy change. Let us denote the entropy change for each J by $\Delta S_{\text{igno}}(J)$. The mean is $\Delta S_{\text{igno}} = \sum_J p_J \Delta S_{\text{igno}}(J)$, and the variance is $V(\Delta S_{\text{igno}}) = \sum_J p_J \Delta S_{\text{igno}}(J)^2 - \Delta S_{\text{igno}}^2$. This can be computed straightforwardly from our expressions for p_J, d_J, and approximated in various limits.

Consider first a high density BEC-limit case with $d = 2$ and $N = 2n \gg 1$ bosons. We have $d_J^B = 2J + 1$, and using the techniques of appendix C.6, $p_J \approx \frac{2J}{n} e^{-J^2/n}$. Then $\Delta S_{\text{igno}} = \sum_J p_J \ln(2J + 1) \approx \frac{1}{2} \ln n + \ln 2 - \frac{\gamma}{2} \approx \frac{1}{2} \ln n + 0.405$. Similarly, we compute $V(\Delta S_{\text{igno}}) = \sum_J p_J [\ln(2J + 1)]^2 \approx \frac{\pi^2}{24} \approx 0.411$. Therefore the mean work dominates its fluctuations (logarithmic versus a constant).

Next, consider the closest analogue for fermions: the case of minimal dimension $d = 2n = 2m$. Recall that $\Delta S_{\mathrm{igno}} \approx \Delta S_{\mathrm{info}} \approx 4n \ln 2$ for large n. Since $p_0 = 1$, work extraction is in fact completely deterministic in this case.

Finally, take the low density limit. As found before, for both bosons and fermions, $\Delta S_{\mathrm{igno}} \approx 2n \ln 2$ – linear in n – and yet we still find a constant $V(\Delta S_{\mathrm{igno}}) \approx \frac{\pi^2}{24}$.

In these macroscopic limits, therefore, work extraction is either fully deterministic or effectively deterministic in that the fluctuations are negligible compared with the mean.

Non-orthogonal spins

The results generalise to the case of partially distinguishable spins – that is, initially with n in spin state $|{\uparrow}\rangle$ on the left and m in state $|{\nearrow}\rangle$ on the right, where

$$|{\nearrow}\rangle = \cos(\theta/2)\,|{\uparrow}\rangle + \sin(\theta/2)\,|{\downarrow}\rangle\,. \tag{4.23}$$

For this, we must be more explicit about the operations permitted by the informed observer. The most general global unitary that does not affect the number of each type of spin is of the form $U = \bigoplus_M U_{xsBW}^{(M)}$, where the block structure refers to subspaces with fixed M as defined by the Schur basis (recalling that the total number of particles is fixed). We find (see appendix C.3 for details) that ΔS_{info} is an average of entropy changes for each value of M. For ΔS_{igno}, all that changes is the probability p_J, now being obtained by an average over Clebsch-Gordan coefficients. Importantly, for both observers, the result is a function of θ only via the probability distribution q_M for the spin value M. In Fig. 4.4, one observes the smooth transition from identical to orthogonal spin states as θ varies from 0 to π.

4.3 Discussion

In contrast to the classical Gibbs paradox setting, we have shown that quantum mechanics permits the extraction of work from apparently indistinguishable gases, without access to the degree of freedom that distinguishes them. It is notable that the lack of information about this "spin" does not in principle impede an experimenter at all in a suitable macroscopic limit with large particle number and low density – the thermodynamical value of the two gases is as great as if they had been fully distinguishable.

The underlying mechanism is a generalisation of the famous Hong-Ou-Mandel (HOM)
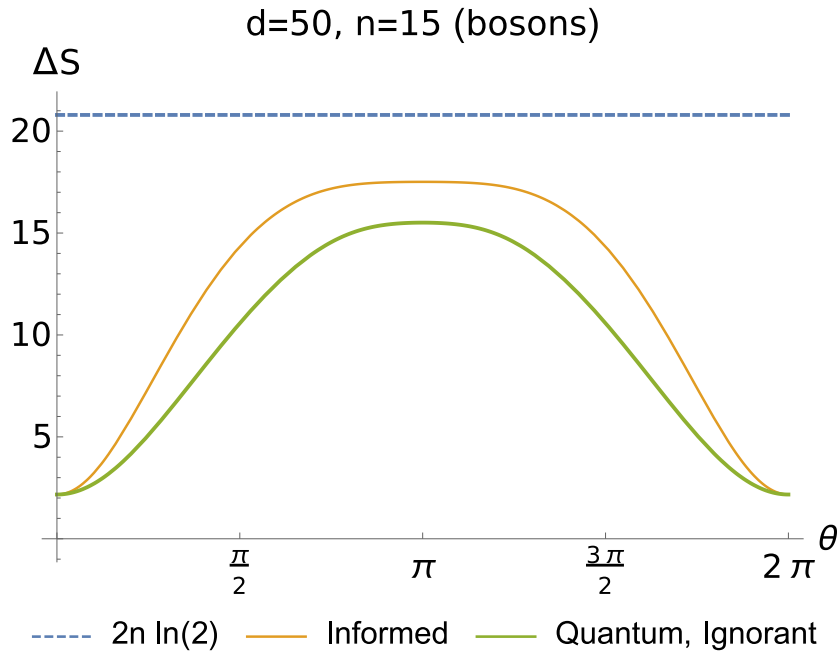
Figure 4.4: **Results for partially distinguishable spins.** Plots of $\Delta S_{\text{info}}, \Delta S_{\text{igno}}$ as a function of orthogonality of the spin states as determined by θ in (4.23). The figure is for a bosonic system with initial numbers of particles on either side of the box $n = m = 15$, and $d = 50$ cells. For comparison, the figure also displays the classical change in entropy, $2n \ln(2)$. Here, the greatest change in entropy occurs when the spin states are orthogonal at $\theta = \pi$.

effect in quantum optics [123, 5, 215]. In this effect, polarisation may play the role of the spin. Then a non-polarising beam splitter plus photon detectors are able to detect whether a pair of incoming photons are similarly polarised. The whole apparatus is polarisation-independent and thus accessible to the ignorant observer. Given this context, it is therefore not necessarily surprising that quantum Gibbs mixing can give different results to the classical case. However, the result of the low density limit is not readily apparent. This limit is reminiscent of the result in quantum reference frame theory [23] that the lack of a shared reference frame presents no obstacle to communication given sufficiently many transmitted copies [22].

Two recent papers [121, 122] have studied Gibbs-type mixing in the context of optome-chanics. There, a massive oscillator playing the role of a work reservoir interacts with the photons via their pressure. There is also a beam splitter between the two sides of the cavity. In Ref. [121], the beam splitter is non-polarising and thus (together with the interaction with the oscillator) accessible to the ignorant observer. The main behaviour there is driven by the HOM effect, which impedes energy transfer to the oscillator. How-ever, this does not contradict our findings: we have shown that an advantage is gained by optimising over all allowed dynamics. It is therefore an interesting question whether such proposals can be modified to see an advantage of the type described here, even if

not optimal. Ref. [122] studies Gibbs mixing as a function of the relative polarisation rotations between left and right. However, this uses a polarising beam splitter, which is only accessible to the informed observer. Therefore the effect described here cannot be seen in such a set-up.

It is important to determine how the thermodynamic enhancements predicted in this paper may have implications for physical systems. Such an investigation should make use of more practical proposals (such as Refs. [121, 122, 172]) to better understand possible realisations of mixing. For example, systems of ultra-cold atoms in optical lattices [136] may provide a suitable platform to experimentally realise the thermodynamic effects predicted in this chapter. The question of the maximal enhancement in the macroscopic limit is particularly compelling given the rapid progress in the manipulation of large quantum systems [87].

Chapter 5

Indistinguishable particle entanglement

5.1   Introduction

In section 1.3 we saw that particles in quantum mechanics have a character quite distinct from those in classical mechanics. In particular how classical indistinguishability comes from limited abilities of the experimenter whereas in the quantum world, two particles of the same type, such as electrons, are fundamentally indistinguishable [84, 228]. This feature applies not only to fundamental particles but is also crucial in describing identical composite particle systems such as Bose-Einstein condensates (BECs) [12]. Notably, exchanging two identical quantum particles results in an overall phase change in the wavefunction: no change for bosons and a minus sign for fermions.

As we saw in previous chapters these exchange statistics require a symmetric or anti-symmetric wavefunction in the first-quantised formalism. For example, let us denote by $|n_0, n_1\rangle$ a state of identical bosons in which n_0, n_1 particles have the internal state $|0\rangle$, $|1\rangle$ respectively. In the first-quantised picture, we represent $|1, 1\rangle$ not as a two-mode state but a symmetric two-particle state

$$\frac{|0\rangle_1|1\rangle_2 + |1\rangle_1|0\rangle_2}{\sqrt{2}}, \tag{5.1}$$

in which we have attached the fictional labels $1, 2$ to the particles. As discussed in section 1.3.1 the state (5.1) is formally *entangled*. However as we examined, it can be argued [78, 95, 94, 225, 224, 67] that this "entanglement" is unphysical – since the particles are identical, the labels $1, 2$ are meaningless as it is impossible to say which particle has

74

which label. Throughout this work we will refer to this manifestation of correlations due to exchange symmetry as *Particle Entanglement* (PE) [1].

A consensus on the nature of this entanglement has so far been out of reach [96, 188, 210, 20, 252, 21, 52, 131, 239, 202, 27, 47, 19, 28, 199, 25, 130]. Some authors view PE as a failure of the mathematical formalism and argue that it should be disregarded in favour of other definitions of identical-particle entanglement [188, 78, 95, 210, 20, 252, 21, 202, 224, 199]. One class of approaches requires talking only about correlations between observables [20, 252, 21, 202, 19, 199]; other authors pursue entirely new definitions of entanglement tailored to the identical-particle setting [96, 188, 95, 78, 210, 94]. Many of these approaches are summarised in a recent review [26].

In order to determine whether there is any meaningful interpretation of PE *per se* we follow the modern resource theoretical approach to entanglement within quantum information theory [128]. Here, entangled states are defined as those which cannot be prepared by two or more separated parties who are unable to send quantum information, and are as such limited to local operations (within their own laboratories) and classical communication – abbreviated as LOCC. Entanglement is then regarded as a *resource* for parties operating under such constraints, and can enable them to perform better at a vast range of tasks including quantum communication [33], computation [156], key distribution [211], and metrology [98], to name a few.

In systems of identical particles, the usable entanglement is that between *modes* [243, 204, 135, 159, 158, 230, 61, 138, 67, 66]. This is because (orthogonal) modes are by definition distinguishable systems and so can be addressed individually. Note that these modes need not be spatially separated; we only require that there exist some degree of freedom (such as momentum or internal spin) via which they can be separately addressed. Mode entanglement is distinct from entanglement between particles. For instance, a single particle existing in a superposition of two locations can be viewed as an entangled state of two spatial modes – but this state clearly contains no PE since there is only one particle. So if mode entanglement is the operationally useful quantity, and is not directly related to PE, why are we interested in the latter? There are strong reasons to believe that PE is a property worth quantifying and may be a resource in certain scenarios. For instance, many-body entangled states of cold atoms, such as spin-squeezed states, can increase precision in metrology thanks to their PE [46, 191, 220, 200, 106, 90].

In order to justify PE as a resource, one needs to provide the appropriate setting – what is the analogue of LOCC for indistinguishable particles? In this chapter, we first answer that question by finding a physically relevant set of quantum operations in which

---

[1]Not to be confused with particle entanglement as named in [230].

PE cannot be created. These operations are constructed from combinations of appending vacuum states, performing passive linear unitaries and making either non-demolition measurements of total particle number, or else arbitrary but destructive measurements. We prove that each of these sets of elements is as general as possible while resulting in a consistent theory. In particular, the set of unitaries is physically motivated as "easy" in many settings, corresponding to beam splitters and phase shifters in optics, and to number-conserving non-interacting hamiltonians in condensed matter systems. These operations, which we call *particle-separable*, define the basis of a *resource theory* for PE. as we saw in section 1.5 such an approach has been widely employed recently to pin down a variety of quantum properties beyond entanglement, such as quantum thermodynamics [41], quantum coherence [216] and asymmetry [22]. With this structure in place, one can begin to rigorously quantify PE and lay the ground for its systematic utilisation in practical tasks.

As a first application, in section 5.4 we consider the metrological value of PE, in the context of sensing rotations around a collective spin observable. It is known that PE can result in a greater Quantum Fisher Information (QFI), a key figure of merit for the estimation precision achievable with a given state [129, 226]. Beyond just acting as a witness for PE, we show that the enhancement in QFI, suitably quantified, is a monotone under particle-separable operations. It thus follows that operations with particle-entangling power are needed to increase the utility of a state for metrology. This provides a fundamental quantitative assessment of the power of PE as a resource in quantum metrology tasks.

In section 5.5 we use our framework to find the complete setting in which PE is a resource for generating useful mode entanglement between parties. This fully generalises earlier observations by Yurke and Stoler [251] and more recently by Killoran et al. [138], the latter providing a starting impetus for this work. Specifically, by "useful" mode entanglement we mean that which is accessible to parties who are constrained not only by LOCC but also by a local particle-number superselection rule [240]. The latter constraint renders superpositions of different particle numbers unobservable, and applies when particle number is conserved and the two parties do not have access to a shared phase reference [23]. In practical terms, this corresponds to the inability to share laser light with a stable relative phase (in optics) or to share a coherently delocalised BEC (with cold atoms). Under this limitation, less entanglement can be utilised [204, 230]. We show that useful entanglement can be generated from an initial state by a particle-separable operation exactly when the initial state contains non-zero PE. Furthermore, we find quantitative relations between the amount of input PE and the output useful entanglement. This shows that PE mirrors other quantum resources which may be similarly "activated" into

useful entanglement [192, 219, 161]. These results provide a full generalisation of the observations in [138]. There, it was found that the Schmidt coefficients of a pure PE state remain invariant during its activation into a useful entangled state under a specific class of unitary operations involving non-polarising beam splitters. Thus we have explored the full resource-theoretic meaning of this activation, for the most general states and operations, and quantified it via large classes of entanglement measures.

Our results have direct applications to real systems of indistinguishable bosons, in particular entangled states of BECs [212, 130]. In section 5.6 we analyse one of a set of recent experimental advances witnessing mode entanglement in BECs [81, 146, 153]. We show that these fit into our framework and implement the above resource conversion. In particular, our results enable for the first time a quantitative determination of the PE content of the states produced in the experiment, based on quantifiers validated within our resource theory framework.

Finally, in section 5.7 we find novel and surprising connections between PE and non-classicality as employed in quantum optics. In that context, classical states are probabilistic mixtures of coherent states [101, 221]. States lying outside this set are non-classical, and are essential in many quantum technological applications [157]. Aided by a recent resource theory formulation of non-classicality [91, 223, 247, 147], several parallels can be formed between the two disparate topics. We find non-classicality to be a necessary but not sufficient prerequisite for PE – however, non-classicality can be "unlocked" by using multiple copies of a state. Thus we have a remarkable link between two uniquely quantum resources.

5.2 Particle identity and superselection rules

We work with bosonic systems, for which m orthogonal modes have associated annihilation and creation operators $a_i, a_i^\dagger$, $i = 0, \ldots, m-1$, satisfying the canonical commutation relations $[a_i, a_j] = 0$, $[a_i, a_j^\dagger] = \delta_{i,j}$. For a particular choice of modes, the second quantised description is given in terms of the occupation numbers n_i of each mode: $|n_0, \ldots, n_{m-1}\rangle \propto (a_{m-1}^\dagger)^{n_{m-1}} \ldots (a_0^\dagger)^{n_0} |0, \ldots, 0\rangle$. All bosonic states then live in the Fock space spanned by such vectors.

In order to make statements about entanglement between particles, it is necessary to ensure that it is even sensible to talk about the particles comprising a state. Such statements are meaningless when a state contains a superposition of different particle numbers. Therefore we permit ourselves only to describe states of definite total particle

number [2] – or probabilistic mixtures of such states [243, 67]. Mathematically, this is described by a particle-number *superselection rule* (SSR), which forces any state ρ under consideration to be block-diagonal with respect to the total number operator $\hat{N}$, also expressed as $[\rho, \hat{N}] = 0$. (We distinguish between the operator $\hat{N}$ and its eigenvalues N.) Similarly, all considered operations $\mathcal{E}$ (i.e., completely positive maps on the set of states) must respect the SSR. This is ensured by taking only *covariant* operations, defined by commutation $[\mathcal{E}, \mathcal{U}_\theta] = 0$ with the phase rotation channel $\mathcal{U}_\theta(\rho) = e^{-i\theta\hat{N}}\rho e^{i\theta\hat{N}}$ for all θ [23]. Equivalently, covariant operations can be performed via a dilation involving an initially number-diagonal environment and a global particle number conserving unitary interaction [137].

Any state of definite particle number $N = \sum_i n_i$ can be written in the first quantised picture, where each particle has an internal state in the single-particle space $\mathcal{H}_1$ of dimension m (so that there is one degree of freedom for each mode). The overall state then lies in the symmetric subspace of the N-system space, denoted by $\mathcal{H}_N = \mathcal{S}[\mathcal{H}_1^{\otimes N}]$. A general mixture of particle numbers $\rho = \sum_N p_N \rho^{(N)}$ can be described as being a state on $\mathcal{S}[\mathcal{H}_1^{\otimes N}]$ with probability p_N. Where necessary in this chapter, we distinguish between the first and second quantised forms of a pure state using the notation $|\psi\rangle^\bullet$ and $|\psi\rangle$ respectively, and similarly $\rho^\bullet$ and ρ for a mixed state.

5.3   PE as a resource

As we saw in section 1.5 a resource theory is defined by two components: the set of *free states* $\mathcal{S}$, which possess no resource, and the set of *free operations* $\mathcal{O}$, which do not add any new resource into the system. (One also tends to think of free operations as possible to perform without any resource, although this interpretation is not always clear.)

The set of free states for PE is straightforward to define. For fixed particle number N, they must be non-entangled (separable) states in the first-quantised picture. Due to symmetry, a pure N-particle free state is thus of the form $|\Psi\rangle^\bullet = |\psi\rangle^{\otimes N}$, also known as a coherent spin state [100, 191]. In second-quantised form, we have $|\Psi\rangle \propto (c_\psi^\dagger)^N |0\rangle$, where $c_\psi^\dagger = \sum_i \psi_i a_i^\dagger$ creates a single particle in an arbitrary mode ψ. A mixed N-particle free state is by definition symmetric and separable – it turns out (see appendix D.1) that this

---

[2]An alternative case can be made: a number superselection rule on operations is often in effect in cold atoms and optics. Then a state ρ_S of system S is operationally equivalent to the dephased state $\Phi_S(\rho_S)$, unless one has access to a phase reference R such as a BEC or laser. But appending an additional system can generally contribute to PE (appendix D.2), so R must be included within the description as a resource. The joint system SR is then described as diagonal in total number.

is equivalent to the form

$$\rho^{\bullet} = \sum_i \lambda_i |\psi_i\rangle\langle\psi_i|^{\otimes N}, \ \lambda_i \geq 0, \ \sum_i \lambda_i = 1. \tag{5.2}$$

Then the full set of free states – which we name *particle-separable* – consists of those $\rho = \sum_N p_N \rho^{(N)}$ such that each of these components in the first-quantised picture is of the form (5.2).

We may then choose as free operations any set that preserves particle-separability. This is required in order to ensure a consistent notion of a resource. There is often tension between the desire for mathematical generality of these operations and wanting them to have a known physical implementation. In our approach, we do not take the largest set of quantum operations preserving particle-separability, but instead construct a physically transparent set from elementary types of operations. We prove that each of these elements is as general as possible.

In the spirit of the Stinespring dilation for quantum operations [180], we construct our free operations out of three basic steps: (i) appending ancilliary modes; (ii) global unitary operations; (iii) projective measurements. We investigate each of these in turn.

(i) Appending ancilliary modes: In mathematical terms, the action of appending to a state ρ another set of modes in a fixed state σ means $\rho \rightarrow \rho \otimes \sigma$ in second quantisation. In order to consider this a free operation, we restrict $\sigma \in \mathcal{S}$. In most resource theories this operation would preserve the set of free states [57]. However, the present theory is unusual in that this generally fails – the simplest example is appending the single-particle state $|1\rangle$ to another copy of itself, as $|1,1\rangle \equiv |1\rangle\,|1\rangle$ is not particle-separable. The reason for this is that appending particles in new modes requires symmetrisation in the first quantised picture, which creates PE. As we show in appendix D.2, the only ancilliary state σ that guarantees preservation of free states is the vacuum.

(ii) Unitaries: The covariance condition for unitaries means that they preserve particle number: $[U, \hat{N}] = 0$. Consider first the component $U^{(N)}$ acting on the N-particle subspace. We see that $U^{(N)}$ preserves $\mathcal{S}$ if and only if it has the first-quantised action $U^{(N)\bullet}|\psi\rangle^{\otimes N} = |\phi\rangle^{\otimes N}$ for every $|\psi\rangle \in \mathcal{H}_1$, where $|\phi\rangle$ can depend on $|\psi\rangle$. Perhaps unsurprisingly, this is equivalent to $U^{(N)\bullet} = u^{\otimes N}$ for any single-particle unitary u, although the argument is not immediate and invokes Wigner's theorem on inner-product-preserving transformations [241] (see appendix D.3). In principle, this u could be different for each number N – however, the introduction of number measurements below implies that we lose no generality by taking a fixed u. Such unitaries have a simple second-quantised description via their action on ladder operators: $U^\dagger a_i^\dagger U = \sum_j u_{ij} a_j^\dagger$, where u_{ij} are the

elements of a unitary matrix. They describe single-particle rotations without interaction, acting identically on all particles, and correspond to passive linear operations in optics, which are easily generated by beam splitters and phase shifters [195].

(iii) Projective measurements: A projective measurement is given by a set of projectors Π_i which are orthogonal and complete: $\Pi_i \Pi_j = \delta_{i,j} \Pi_i$, $\sum_i \Pi_i = \mathbb{1}$. As for unitary operations, these must adhere to the SSR, $[\Pi_i, \hat{N}] = 0$, and preserve the set of particle-separable states, $\Pi_i^{\bullet(N)} |\psi\rangle^{\otimes N} \propto |\phi\rangle^{\otimes N}$. However, we find that these conditions are only met by a measurement of total particle number (see appendix D.4). In order to enlarge the set of available measurements, we allow *destructive measurements*, in which the measured modes are subsequently discarded. In appendix D.4 we demonstrate that this relaxation allows any measurement adhering to the SSR to be performed on the system without introducing PE. Such destructive measurements correspond to the majority of experimental photon- and atom-counting techniques.

The set $\mathcal{O}$ of *particle-separable operations* is defined as all possible protocols which result from combinations of the above elements, including possible conditioning of future operations on the results of measurement outcomes. We also allow for the use of classical randomness and coarse-graining – i.e., forgetting measurement outcomes. Mathematically, an element in $\mathcal{O}$ is represented as a quantum instrument, which is a set of CP maps $\mathcal{E}_i$ where each i labels a single (possibly coarse-grained) measurement outcome and the sum $\sum_i \mathcal{E}_i$ is deterministic (trace-preserving). Note that an instrument can equivalently be represented as a deterministic channel $\mathcal{F}(\rho) = \sum_i \mathcal{E}_i(\rho) \otimes |i\rangle\langle i|_X$, where the outcome is stored in a classical system X [6].

With this structure in place, we can now move naturally to define measures M_{PE} of PE. As is standard in quantum resource theories [57], we require that any measure of PE fulfills the following three conditions. *Condition (i)*–It must not detect PE when there is none, meaning $M_{\mathrm{PE}}(\rho) = 0$ for all $\rho \in \mathcal{S}$ (and optionally the converse may be required). *Condition (ii)*–M_{PE} must be a *monotone*, i.e. cannot increase under the action of any particle-separable operation. This reflects the idea that particle-separable operations cannot inject additional PE into the system. Monotonicity can be stated either deterministically, $M_{\mathrm{PE}}(\rho) \geq M_{\mathrm{PE}}(\mathcal{E}[\rho])$ for any channel $\mathcal{E} \in \mathcal{O}$, or probabilistically, $M_{\mathrm{PE}}(\rho) \geq \sum_i p_i M_{\mathrm{PE}}(\rho_i)$ for an instrument $\{\mathcal{E}_i\}$ in $\mathcal{O}$ with outcomes $p_i \rho_i = \mathcal{E}_i(\rho)$. *Condition (iii)*–Convexity, i.e., being non-increasing under probabilistically mixing different states, $\sum_i p_i M_{\mathrm{PE}}(\rho_i) \geq M_{\mathrm{PE}}(\sum_i p_i \rho_i)$.

A straightforward class of PE measures are given by the minimal distance between a

state and the set of particle-separable states:

$$M_{\text{PE}}^D(\rho) := \min_{\sigma \in \mathcal{S}} D(\rho, \sigma), \tag{5.3}$$

where D is any suitable measure of distinguishability between two quantum states. Conditions (i,iii) and the deterministic version of (ii) are met whenever D is contractive under quantum channels (so that $D(\mathcal{E}(\rho), \mathcal{E}(\sigma)) \leq D(\rho, \sigma)$ for any channel $\mathcal{E}$) and jointly convex in its arguments; other properties may guarantee ensemble monotonicity (ii) (see appendix D.5 and Ref. [57]).

5.4 Quantifying metrological power of PE

Now that we have determined the set of protocols under which PE may abstractly be considered a resource, we are in a position to demonstrate concrete tasks in which it is useful. In this section, we use our resource theory to demonstrate a quantitative connection between PE and quantum metrology. A typical metrological setting involves a parameter θ encoded into a system, such that the experimenter is given one of a parameterised family of states ρ_θ, and the task is to estimate θ via measurements. Here, we focus on the case of unitary encoding, whereby an initial state ρ evolves under a given Hamiltonian H, so that $\rho_\theta = e^{-i\theta H} \rho e^{i\theta H}$. An important figure of merit is the quantum Fisher information (QFI) $\mathcal{F}(\rho, H) := -4\partial_\theta^2 \text{Fid}(\rho, \rho_\theta)|_{\theta=0}$, where $\text{Fid}(\rho, \sigma) = \text{Tr}\sqrt{\sqrt{\rho}\sigma\sqrt{\rho}}$ is the fidelity between two states. The QFI can be thought of as a measure of speed of evolution for ρ_θ under the dynamics generated by H. Its importance for metrology is given by the (quantum) Cramér-Rao bound, which says that the uncertainty $\Delta\theta$ in estimating θ is lower-bounded by $(\Delta\theta)^2 \geq 1/(n\mathcal{F}(\rho, H))$ with n copies of ρ_θ provided [185].

PE is known to be a necessary resource for a quantum-enhanced metrology [46, 191]. For N qubits, one can define total spin components $\boldsymbol{S}^\alpha := \sum_{i=1}^N \sigma_i^\alpha/\sqrt{N}$, $\alpha = x, y, z$, where σ_i^α is a Pauli matrix acting on the ith particle; the spin in any direction $\boldsymbol{n} = (n^x, n^y, n^z)$, with $|\boldsymbol{n}| = 1$, is denoted as $\boldsymbol{n}\cdot\boldsymbol{S}$. Then, for any particle-separable state, we have $\mathcal{F}(\rho, \boldsymbol{n}\cdot\boldsymbol{S}) \leq 1$ [226, 129]. Exceeding this bound witnesses PE, with the maximum possible QFI being N. A tighter bound, applicable to any Hamiltonian of the form $H = \sum_{i=1}^N h_i/\sqrt{N}$, was more recently proven [93]:

$$\rho \in \mathcal{S} \Rightarrow \mathcal{F}(\rho, H) \leq 4\sum_{i=1}^N V\left(\rho, \frac{h_i}{\sqrt{N}}\right) = 4V(\rho, h_1). \tag{5.4}$$

Based on this inequality, we define the following quantity as the amount by which the

QFI exceeds the limit for particle-separable states:

$$M_{\mathrm{PE}}^{\mathcal{F}}(\rho) := \max_{h:\|h\|=1} \left[\mathcal{F}(\rho, H) - 4V(\rho, h)\right]^+ , \tag{5.5}$$

where $H = \bigoplus_N H^{(N)}$, $H^{(N)\bullet} = \sum_{i=1}^N h_i/\sqrt{N}$, $[x]^+ = \max\{x, 0\}$ denotes the positive part of x, and the maximisation is performed over all single-particle observables h with unit operator norm. The expectation value of a single-particle operator h in a number-varying state $\rho = \sum_N p_N \rho^{(N)}$ is defined as

$$\langle h \rangle_\rho := \sum_N p_N \mathrm{Tr}\left[\rho^{(N)\bullet} h_1\right] = \sum_N p_N \frac{1}{N} \mathrm{Tr}\left[\rho^{(N)\bullet} \sum_{i=1}^N h_i\right], \tag{5.6}$$

so that $V(\rho, h) := \langle h^2 \rangle_\rho - \langle h \rangle_\rho^2$.

We can also extend the measure to include settings where one records measurement outcomes in a classical memory M. In this case, a state is in "quantum-classical" form $\rho_{SM} = \sum_m p_m \rho_{S|m} \otimes |m\rangle\langle m|_M$, where p_m is the probability of outcome m, $\rho_{S|m}$ the corresponding conditional state of the system S, and the states $\{|m\rangle\}$ form an orthonormal basis for the memory M. For such a state, the observable h is understood to only act on S and not on the memory M, i.e.,

$$M_{\mathrm{PE}}^{\mathcal{F}}(\rho_{SM}) := \max_{h_S:\|h_S\|=1} \left[\mathcal{F}(\rho_{SM}, H_S) - 4V(\rho_{SM}, h_S)\right]^+ . \tag{5.7}$$

As a consequence of this definition, the QFI part can be expressed as an average over measurement outcomes, $\sum_m p_m \mathcal{F}(\rho_{S|m}, H_S)$, while the variance part is calculated for the whole ensemble ρ_{SM}.

Remarkably, we find that $M_{\mathrm{PE}}^{\mathcal{F}}$ is not only a witness of PE, but also a monotone under particle-separable operations (without feed-forward):

Theorem 4. $M_{PE}^{\mathcal{F}}$ *is convex and satisfies* $M_{PE}^{\mathcal{F}}(\rho) = 0 \, \forall \rho \in \mathcal{S}$. *Moreover, let* $\mathcal{E}_{S \to SM} \in \mathcal{O}$ *contain a single measurement round, such that no conditional operations are performed after the measurement. We may write* $\mathcal{E}_{S \to SM}(\rho_S) = \sum_m \mathcal{E}^m(\rho_S) \otimes |m\rangle\langle m|_M$, *where* $\mathcal{E}^m$ *is the operation applied to* ρ_S *conditioned on outcome* m. *Then*

$$M_{PE}^{\mathcal{F}}(\rho_S) \geq M_{PE}^{\mathcal{F}}(\mathcal{E}_{S \to SM}[\rho]). \tag{5.8}$$

The proof is presented in appendix D.6. Note that $M_{\mathrm{PE}}^{\mathcal{F}}$ may vanish for some particle-entangled states – however, for pure states, it does faithfully detect PE [93]. The monotonicity result demonstrates that, beyond being a witness, $M_{\mathrm{PE}}^{\mathcal{F}}$ captures the ordering of

particle-entangled states under the free operations in the resource theory developed in this paper. From a practical perspective, this shows the limitations on particle-separable operations for enhancing the utility of a state for metrology, and ultimately provides an original and operationally motivated tool to quantify PE by means of its metrological value, in addition to the distance-based measures presented earlier.

A simplification is possible in the special case of two modes (i.e., when the particles are qubits). Given $\|h\| = 1$, without loss of generality we can write $h = |0\rangle\langle0| + \lambda|1\rangle\langle1|$ in some basis, where $|\lambda| \leq 1$. Since the QFI and variance are invariant under constant shifts of the observable, we can shift h to $h - (\frac{1+\lambda}{2})I = (\frac{1-\lambda}{2})\sigma^z$, thus getting

$$[\mathcal{F}(\rho, H) - 4V(\rho, h)]^+ = \left(\frac{1-\lambda}{2}\right)^2 [\mathcal{F}(\rho, Z) - 4V(\rho, \sigma^z)]^+$$
$$\leq [\mathcal{F}(\rho, Z) - 4V(\rho, \sigma^z)]^+, \tag{5.9}$$

where $Z^{(N)\bullet} = \sum_{i=1}^{N} \sigma_i^z/\sqrt{N}$. Equality is obtained for $\lambda = -1$, i.e., $h = \sigma^z$. Hence, in this case, the only remaining degree of freedom is the eigenbasis of h, which can be translated into a spin direction $\boldsymbol{n}$:

$$\dim \mathcal{H}_1 = 2 \Rightarrow M_{\text{PE}}^{\mathcal{F}}(\rho) = \max_{\boldsymbol{n}:|\boldsymbol{n}|=1} [\mathcal{F}(\rho, \boldsymbol{n}\cdot\boldsymbol{S}) - 4V(\rho, \boldsymbol{n}\cdot\boldsymbol{\sigma})]^+. \tag{5.10}$$

Note how, in addition to generalising (and tightening) the QFI witnesses proposed in Refs. [226, 129], our measure $M_{\text{PE}}^{\mathcal{F}}$ differentiates itself by explicitly including the variance of the single-particle observable, rather than being used to bound the measure. The importance of its inclusion is apparent in the proof of Theorem 4, specifically in order to show that $M_{\text{PE}}^{\mathcal{F}}$ is invariant under the addition of vacuum modes. When new modes are included, the set of possible h observables increases, allowing for a greater possible QFI – we may have $\max_{h'} \mathcal{F}(\rho_S \otimes |0\rangle\langle0|_A, H') > \max_h \mathcal{F}(\rho_S, H)$. The variance component nontrivially compensates for this effect.

5.5 Activating PE

Here, we describe another important task for which the utility of PE as a resource is manifest. The original seeds of the *activation* protocol that we analyse here are in work by Yurke and Stoler, who noted that two particles produced from separated, independent sources can in fact be used to violate a Bell inequality [251]. The protocol that we present is a direct application of our resource theoretic formulation and constitutes a full generalisation of [138].
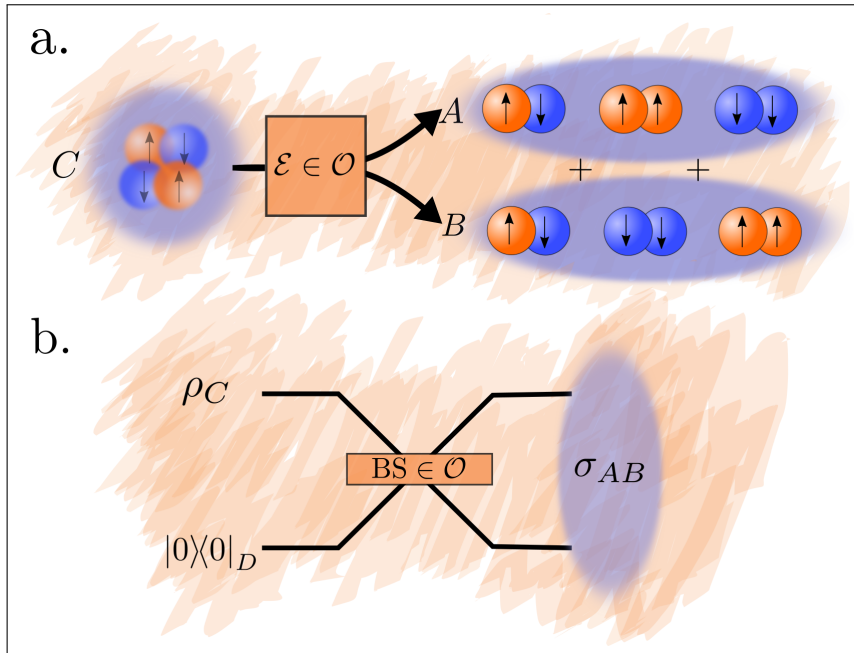
Figure 5.1:   **a.** Conversion protocol between PE and SSR-entanglement via the quantum operation $\mathcal{E} \in \mathcal{O}$. The operation $\mathcal{E}$ converts a system of identical particles with PE into a bipartite state, whose SSR-entanglement can be extracted and utilised in quantum information tasks. The above diagram depicts the transformation $|2,2\rangle_C \xrightarrow{\mathcal{E} \in \mathcal{O}}$ $(|1,1\rangle_A |1,1\rangle_B + |2,0\rangle_A |0,2\rangle_B + |0,2\rangle_A |2,0\rangle_B)$, having post-selected $N_A = N_B = 2$.
b. An example of a particle-separable operation is the action of a beam-splitter with a vacuum, which can be used to activate the PE present in the state ρ_C.

Consider two separated parties, A and B, who want to perform some joint quantum information protocol but are constrained to classical communication and additionally lack a shared phase reference (conjugate to the number observable $\hat{N}_A$ or $\hat{N}_B$). A phase reference would be provided by a shared state containing coherence with respect to the local number observable $\hat{N}_A$ (or $\hat{N}_B$). In optics, a typical example is a laser coherently split into modes held by each party, maintaining a fixed phase relationship. The analogue in cold atoms is a coherently distributed BEC. Extensive discussions of the relationship between SSRs and phase references can be found in Refs. [67, 23].

While each party may be unconstrained in their local operations, without sharing a phase reference, the amount of entanglement accessible to them is reduced by the application of an effective local SSR [23]. This SSR corresponds to both local particle numbers $\hat{N}_A$ and $\hat{N}_B$. A third party C is tasked with providing A and B with a shared entangled state that they can use. To accomplish this, C has an initial resource state ρ_C of m modes and can process it using any particle-separable operation $\mathcal{E}$ before distributing m_A and m_B modes to each of A and B. (Recall that the operation $\mathcal{E}$ may introduce new vacuum modes and trace out some modes; see Fig. 5.1). The question is: how much useful entanglement can be extracted in this way from ρ_C?

Let $\sigma_{AB} = \mathcal{E}(\rho_C)$ be the output state sent to A and B, where $\mathcal{E} \in \mathcal{O}$ is the distribution operation performed by C. (Without loss of generality, using classical flags, we can take this to be deterministic.) Due to the local SSR, from the perspective of A and B, this state is operationally as useful as the state $\Phi_A \otimes \Phi_B(\sigma_{AB})$ [205], where Φ_S is the dephasing channel local to subsystem S, removing quantum coherences between states of differing local number $\hat{N}_S$ [3].

For any measure E of bipartite entanglement, we can then define the corresponding measure of entanglement accessible to A and B [23]:

$$E_{\mathrm{SSR}}(\sigma_{AB}) := E\left(\Phi_A \otimes \Phi_B(\sigma_{AB})\right) \leq E(\sigma_{AB}). \tag{5.11}$$

We say that a state σ_{AB} is *SSR-separable* whenever it has vanishing accessible entanglement – i.e., when $\Phi_A \otimes \Phi_B(\sigma_{AB})$ is separable – and *SSR-entangled* otherwise. The inequality in (5.11) follows from the fact that $\Phi_A \otimes \Phi_B$ is a local operation – the local SSR generally reduces the amount of accessible entanglement. The aspect of the entanglement in σ_{AB} that is *inaccessible*, sometimes referred to as "fluffy bunny entanglement" [244], is connected with superpositions of local number. Note that Wiseman and Vaccaro [243] proposed the same class of measures (5.11) and found such SSR-entanglement to require non-zero PE in the case of two particles.

We prove that PE in the initial state ρ_C is precisely the resource enabling the distribution of SSR-entanglement. Our first result is that the mapping between the two types of entanglement is faithful, in that SSR-entanglement can be extracted exactly when there is nonzero PE (see appendix D.8 for the proof):

Theorem 5. *There exists an activation operation $\mathcal{E}_{C \to AB} \in \mathcal{O}$ creating an SSR-entangled state σ_{AB} from ρ_C if and only if $\rho_C \notin \mathcal{S}$.*

Moreover, almost any operation of the following type is sufficient to activate PE into non-zero SSR-entanglement: for each mode i in C, attach a new mode in the vacuum state, and perform a global passive-linear unitary coupling the modes (as in Fig. 5.1b). We say "almost all" because the unitary must not be trivial by failing to couple some of the modes. Ref. [138] examined activation for a specific class of unitary interactions, namely a set of beam-splitters with identical transmission coefficients. However, we see that a much more general statement is possible, expanding the scope to all particle-separable operations.

---

[3]This may be written equivalently as a phase average $\Phi_S(\rho) = \int_0^{2\pi} \mathrm{d}\theta\, e^{-i\theta \hat{N}_S} \rho e^{i\theta \hat{N}_S} / 2\pi$ or as a "measure-and-forget" operation of the local number: $\Phi_S(\rho) = \sum_n P_{n,S} \rho P_{n,S}$, where $P_{n,S}$ is the projector onto the subspace of n particles in S.

Beyond the faithful mapping between nonzero resources, we now quantitatively relate the input and output forms of entanglement. One approach uses measures of both PE and SSR-entanglement constructed in the same way. Recall the distance-based measure of PE M_{PE}^D; by the same recipe, one can construct a measure of SSR-entanglement (see appendix D.8):

$$
\begin{aligned}
E_{\text{SSR}}^D(\rho_{AB}) &= E^D(\Phi_A \otimes \Phi_B[\rho_{AB}]) \\
&:= \min_{\sigma_{AB} \in \text{ sep.}} D(\Phi_A \otimes \Phi_B[\rho_{AB}], \sigma_{AB}).
\end{aligned}
\tag{5.12}
$$

As shown in appendix D.7, when ρ respects the local SSR, the minimisation can be equivalently performed over the smaller set of σ_{AB} being separable and respecting the local SSR. Using this, we have:

Theorem 6. *For any activation $\mathcal{E}_{C \to AB} \in \mathcal{O}$, $E_{SSR}^D(\mathcal{E}_{C \to AB}[\rho_C]) \leq M_{PE}^D(\rho_C)$.*

This shows that the amount of accessible entanglement extracted never exceeds the initial amount of PE. Note, however, a subtlety: in general, this inequality is strict (apart from when both sides are zero), due to a necessary reduction in entanglement after applying the dephasing operation $\Phi_A \otimes \Phi_B$ and removing the "fluffy bunny entanglement".

Alternatively, we can take any measure of SSR-entanglement and use it to construct a new measure of PE. This is given by the maximal amount of SSR-entanglement which can be created from a certain initial state:

Theorem 7. *For any (convex) entanglement measure E, the quantity defined as*

$$
M_{PE}^E(\rho) := \sup_{\mathcal{E}_{C \to AB} \in \mathcal{O}} E_{SSR}\left(\mathcal{E}_{C \to AB}[\rho_C]\right)
\tag{5.13}
$$

is a (convex) measure of PE.

In other words, for any entanglement measure E, the corresponding quantity M_{PE}^E satisfies criteria (i-iii). Theorem 7 gives a precise quantitative version of the statement that PE is the resource for producing SSR-entanglement.

5.6   Experimentally measuring PE

In this section we demonstrate that our resource theory for describing PE and its activation encompasses recent experimental investigations [81, 146, 153] converting PE into useful mode entanglement. This enables us to promptly analyse the experimental data

from [81] in order to extract a lower bound to a measure of PE. To the best of our knowledge, this constitutes the first instance of quantitative estimation of PE in an experiment.

The experimental method is as follows – see [81] for more details. The BEC is initialised in a spin-squeezed state [213], which possesses PE. This state being prepared through atomic collisions with a state-dependant potential on an atom chip, as described by [182]. The BEC is then released from its trap and allowed to expand during a 2.2ms time of flight. During the expansion, the effect of interactions between atoms on their spin state is negligible such that this step can be regarded as a beam-splitter operation. Namely that an individual atoms has a 50:50 chance of being on either side of the split BEC post expansion. These dynamics therefore fall within our set of particle-separable operations [4]. In order to measure the spin components of the two spatially separated clouds we set the spin axis by applying a Rabi rotation pulse to the entire atomic cloud then by illuminating the atomic cloud with resonant laser pulses we record two high resolution atomic absorption images to determine the atomic density distribution of the two atomic states. These imaging pulses both project the atom into a well defined spin state and localize its position. It should be noted that the measurement of spin components of the spatially separated regions adheres to the local SSR [5]. This experimental procedure is then repeated several thousand times, where the spin measurement direction is alternated between x, y and z.

The correlations between the two spatial regions are held in the spin components of the condensate atoms. In particular the z-component of the spin in regions A, B is defined as $\hat{S}_z^{(A,B)} := \frac{1}{2\eta_{\text{eff}}^{(A,B)}} \left(\hat{N}_1^{(A,B)} - \hat{N}_2^{(A,B)} \right)$ where $1, 2$ correspond to the internal degree of freedom of the atom and $\eta_{\text{eff}}^{(A,B)}$ accounts for finite spatial resolution in the detection of the BEC. Other spin components, e.g. $\hat{S}_x^{(A,B)}$ and $\hat{S}_y^{(A,B)}$, can be measured by applying appropriate spin rotations before detection, these local rotations also being allowed within SSR constraints.

In Ref. [81] the authors showed how these local spin measurements can violate the inequality [99]

$$\frac{4\text{Var}\left(g_z\hat{S}_z^A + \hat{S}_z^B\right)\text{Var}\left(g_y\hat{S}_y^A + \hat{S}_y^B\right)}{\left(|g_zg_y|\left|\left\langle\hat{S}_x^A\right\rangle\right| + \left|\left\langle\hat{S}_x^B\right\rangle\right|\right)^2} \geq 1 \,, \tag{5.14}$$

in terms of variances and average values of spin observables. The condition (5.14) is

---

[4]The interaction of ultracold ^{87}Rb atoms depends only very weakly on their spin state. During the expansion of the BEC, the interactions therefore do not affect the spin state and are furthermore quickly rendered small due to the decreasing density [51].

[5]Due to technical limitations a fraction of the atomic spins in a gap between the two regions is discarded in the measurement process.
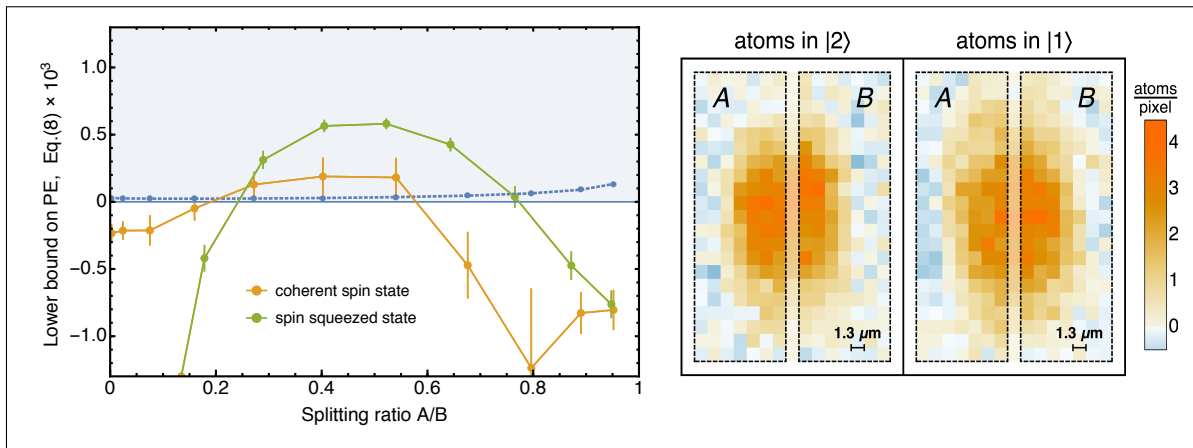
Figure 5.2: Based on the measurements [81] we are able to extract the lower bound given by the right-hand side of (5.15), on the PE measure $M_{\mathrm{PE}}^{\mathrm{Tr}}$. The two sets of points correspond to initialising the BEC either in a spin squeezed state (green), where Particle Entanglement is present, or in a coherent spin state (orange), which is particle-separable. Along the horizontal axis we vary the relative size of the two regions A and B from which we extract the spin values as explained in [81]. In the experiment, technical limitations in the resolution of assigning the atomic spins to the regions can lead to classical correlations, resulting in apparent entanglement. We give an upper bound for these correlations as the blue dashed line. For intermediate splitting ratios we find significant entanglement in the case of the spin squeezed state while the coherent spin state remains compatible with no particle entanglement within experimental error. On the right we show single-shot absorption images of the atomic densities for the two internal degrees of freedom, with an example of regions A and B used to define the collective spins $\hat{S}^A$ and $\hat{S}^B$ entering in (5.15).

satisfied by all separable states and for any real constants $g_{y,z}$, therefore certifying entanglement between system A and B whenever a violation is measured. In appendix D.9, we linearise (5.14) and use Theorem 7 to derive a lower bound on a measure of PE:

$$M_{\mathrm{PE}}^{\mathrm{Tr}}(\rho) \geq \frac{-1}{\mathcal{N}} \left[\mathrm{Var}\left(g_z \hat{S}_z^A + \hat{S}_z^B \right)_\rho + \mathrm{Var}\left(g_y \hat{S}_y^A + \hat{S}_y^B \right)_\rho \right.$$

$$\left. - \left\langle |g_z g_y| \, \hat{S}_x^A + \hat{S}_x^B \right\rangle_\rho \right],$$

$$\mathcal{N} := \frac{1}{4} \left(\frac{|g_z| N_1^A}{\eta_{\mathrm{eff}}^A} + \frac{N_1^B}{\eta_{\mathrm{eff}}^B} \right)^2 + \frac{1}{4} \left(\frac{|g_y| N_1^A}{\eta_{\mathrm{eff}}^A} + \frac{N_1^B}{\eta_{\mathrm{eff}}^B} \right)^2$$

$$+ \left(\frac{|g_z g_y| N_1^A}{\eta_{\mathrm{eff}}^A} + \frac{N_1^B}{\eta_{\mathrm{eff}}^B} \right), \tag{5.15}$$

where $M_{\mathrm{PE}}^{\mathrm{Tr}}$ is defined according to (5.3) with the trace distance $D_{\mathrm{Tr}}(\rho, \sigma) := \frac{1}{2} \mathrm{Tr} \, |\rho - \sigma|$. We show an evaluation of this bound using experimental results in Fig. 5.2. The parameters $g_{y,z}$ are optimised numerically so that the left-hand side of (5.14) is minimised, as this expression is more robust than (5.15) against experimental noise. This plot clearly shows a positive amount of PE has been activated from a spin squeezed BEC and none

from a coherent spin BEC state, as predicted from our theory.

The case study presented in this section reveals how our resource theoretic character-isation of PE unlocks useful quantitative tools that can be readily employed by the cold atoms community to benchmark present and future experiments, including demonstrations of entanglement production and manipulation, sensing and metrology tasks, and other quantum technology protocols empowered by PE.

5.7 Connections to non-classicality

While coherent spin states are considered classical in cold atoms settings with fixed par-ticle number, continuous-variable coherent states in quantum optics provide the model of classical light. Non-classical states display features such as photon anti-bunching, sub-poissonian statistics and squeezing [92], and form the basis of many quantum techno-logical applications [157] As has been recently appreciated, [91, 247, 147] non-classicality can also be quantified with its own resource theory. In this section we demonstrate some remarkable connections between the resources theories for PE and non-classicality.

Recall that a single-mode coherent state $|\alpha\rangle$ is an eigenstate of the annihilation operator: $a|\alpha\rangle = \alpha|\alpha\rangle$, and a multi-mode coherent state may be written as $|\boldsymbol{\alpha}\rangle := |\alpha_1\rangle \ldots |\alpha_m\rangle$, where $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_m) \in \mathbb{C}^m$. A state is called *classical* if it can be written as a probabilistic mixture of coherent states:

$$\rho = \int \mathrm{d}^{2m}\boldsymbol{\alpha}\, P(\boldsymbol{\alpha})\, |\boldsymbol{\alpha}\rangle\langle\boldsymbol{\alpha}|, \quad P(\boldsymbol{\alpha}) \geq 0. \tag{5.16}$$

Due to the SSR employed here, we restrict to number-diagonal (ND) classical states – i.e, those satisfying $[\rho, \hat{N}] = 0$.

The operationally motivated free operations for non-classicality, presented in Ref. [247], are very close to particle-separable operations. The only differences are that (i) rather than only the vacuum, any classical state may be prepared for free in a new mode and (ii) non-destructive measurements of total particle number can create non-classicality. Moreover, there is an entirely analogous protocol activating non-classicality into mode entanglement [140, 235, 13] (which in fact extends to more general notions of non-classicality [139]). Whereas PE can be activated under particle-separable operations into SSR-entanglement, nonclassicality activates into entanglement accessible without local SSR constraints – equivalently, entanglement which can be accessed when a shared phase reference is avail-able.

This observation immediately implies a relation between the free states of the two resource theories: *all ND classical states are particle-separable.* This follows from the fact that a classical state is always activated onto a separable state, which is always also SSR-separable, implying via Theorem 5 that the input is particle-separable. In fact, this can be shown by a more direct argument, with details in appendix D.10. Essentially, any multi-mode coherent state $|\boldsymbol{\alpha}\rangle$ can be regarded as a single-mode state – for any choice of mode decomposition, there is always a passive linear unitary U such that $U|\boldsymbol{\alpha}\rangle = |\bar{\alpha}\rangle|0\ldots0\rangle$, where $|\bar{\alpha}|^2 = \sum_{i=1}^{m}|\alpha_i|^2$. So any classical state is a probabilistic mixture of terms in which all particles occupy the same mode.

Evidently, ND classical states form a strict subset of particle-separable states. Consequently, we may say that nonclassicality is lower-bounded by PE in the sense that, for any distance measure of nonclassicality M_{NC}^D constructed in the manner of (5.3), the inequality $M_{\mathrm{NC}}^D \geq M_{\mathrm{PE}}^D$ holds.

What distinguishes the two sets of free states? As noted earlier, a striking property of PE is that multiple copies of a free state ρ do not in general jointly form a free state. Viewed through the activation protocol, this is equivalent to saying that two copies of an SSR-separable state may be SSR-entangled. This is possible because of the way the SSR behaves for multiple copies of a system [230, 204]. If A and B share two pairs of entangled systems, (A_1, B_1) and (A_2, B_2), then the particle number local to A is $\hat{N}_A = \hat{N}_{A_1} + \hat{N}_{A_2}$ and similarly for B. The local SSR is applied by $\Phi_A \otimes \Phi_B \neq \Phi_{A_1} \otimes \Phi_{A_2} \otimes \Phi_{B_1} \otimes \Phi_{B_2}$. The lack of factorisation is due to degeneracy in the eigenvalues of $\hat{N}_A, \hat{N}_B$. For example, $(|0\rangle_A|1\rangle_B + |1\rangle_A|0\rangle_B)/\sqrt{2}$ is entangled but SSR-separable; the two copy state

$$
\begin{aligned}
&\frac{1}{2}\left(|0\rangle_A|1\rangle_B + |1\rangle_A|0\rangle_B\right)^{\otimes 2} = \\
&\frac{1}{2}\left(|00\rangle_A|11\rangle_B + |01\rangle_A|10\rangle_B + |10\rangle_A|01\rangle_B + |11\rangle_A|00\rangle_B\right)
\end{aligned}
\tag{5.17}
$$

is SSR-entangled thanks to correlations in the block $N_A = N_B = 1$. This phenomenon is closely related to *work-locking* in quantum thermodynamics, whereby coherence in one copy of a state is useless for work extraction but becomes usable in two copies [160].

A tensor product of two classical states is always classical, hence multiple copies of an ND classical state always have zero PE. Are these the only states with this property? We first consider *number-bounded* states: those for which the expansion $\sum_N p_N \rho^{(N)}$ terminates at a finite maximum. In this case, the resource content of two copies is sufficient to distinguish the classical subset of particle-separable states (note that all classical states apart from the vacuum are necessarily unbounded in number):

Theorem 8. *Two copies $\rho^{\otimes 2}$ of a number-bounded state ρ are particle-separable if and*

only if ρ is the vacuum.

(See the proof in appendix D.10.) In the general unbounded case, let us first take pseudo-pure states, by which we mean those obtained by applying the SSR to a pure state: $\rho = \Phi(|\psi\rangle\langle\psi|)$. It is known that in the limit $k \to \infty$ of many copies $|\psi\rangle^{\otimes k}$ of a pure entangled state, the SSR is effectively lifted in that the full entanglement entropy is distillable [204]. One may then argue from the activation protocol as follows: a non-classical state at the input results in entanglement at the output; many copies of this state must therefore result in an SSR-entangled state. Hence any non-classical pseudo-pure state must fail to be particle-separable with sufficiently many copies. An even stronger statement is in fact possible:

Theorem 9. *Two copies $\Phi(|\psi\rangle\langle\psi|)^{\otimes 2}$ of a pseudo-pure state are particle-separable if and only if $|\psi\rangle$ is classical.*

Therefore we see that non-classicality of any pseudo-pure state, even if particle-separable, can always be unlocked into non-zero PE by taking only two copies.

Finally, we prove the strongest possible connection between particle-separable and classical states, which concerns the case of arbitrarily many copies. The only assumption here is of a finite mean particle number (and, as usual, $\rho = \Phi(\rho)$).

Theorem 10. *Let ρ have finite mean particle number, $\mathrm{Tr}\left[\rho\hat{N}\right] < \infty$, and suppose that $\rho^{\otimes k}$ is particle-separable for some k. Then the trace-distance non-classicality of ρ is bounded by*

$$M_{\mathrm{NC}}^{\mathrm{Tr}}(\rho) \leq \frac{1}{k}. \tag{5.18}$$

Consequently, $\rho^{\otimes k}$ is particle-separable for all k if and only if ρ is classical.

The importance of this result is the realisation that every (finite mean number) non-classical state has the potential to contain particle entanglement, and thus all of the associated resource value, once sufficiently many copies are taken. The proof (in appendix D.10) follows from a novel de Finetti-type theorem, which may be of independent interest.

5.8   Discussion

We have shown that entanglement between identical particles, despite its seemingly fictitious nature, is described by a consistent resource theory whose free operations are

implementable in a wide range of physical systems. Far from just an abstract quantity, this particle entanglement can be quantified by virtue of the advantage it yields for quantum metrology, and can be activated, via the same types of free operations, into directly accessible mode entanglement. This occurs in a setting where phase references are not easily shared between separated parties, enforcing a local SSR.

While we have found the most general form that such an activation may take, some important questions remain open. Theorem 7 expresses the maximum activated SSR-entanglement from a given state as a measure of PE – however, because of our construction we can raise the following question: What is the optimal operation to activate this entanglement? This may depend on the measure being employed, but it is plausible that such an optimal operation should be unitary; Lemma 5 in appendix D.8 proves a simplification from the full space of passive linear unitaries down to only one real parameter per mode, making the optimisation feasible.

Our formulation reveals PE as fundamentally connected not only to entanglement under SSRs, but also to continuous variable non-classicality. In particular, we have shown that SSR-compliant classical states possess no PE. Consequently, PE is a stronger (rarer) resource than non-classicality. Nevertheless, by utilising multiple copies of a state, one may unlock its non-classicality into PE. This unlocking is possible with two copies of any pure non-classical state; in general, non-classicality always results in PE after taking sufficiently many copies. Hence, in a sense, non-classicality emerges as a many-copy limit of PE. It is worth exploring other quantitative ways in which this limit may manifest itself.

It is also worth noting some similarity with other resource theories. For instance, the structure of particle-separable operations bears some resemblance to "strictly incoherent operations", a set of free operations for quantum coherence [248]. Without measurements, particle-separable operations coincide with the zero-temperature limit of a recent treatment of continuous-variable thermodynamics [174] (see also the related approach [208]). One could therefore explore thermodynamical consequences of PE in future work.

Finally, we would like to motivate the wider theoretical and experimental applicability of our framework for PE. In addition to describing the metrological power and the activation of entanglement from a BEC, the framework applies to any system of identical bosons, opening up the possibility of investigating PE beyond BECs and optics, to other condensed matter systems in which entanglement is of interest, such as superfluid Helium [119].

A study of PE in fermionic systems could also be pursued, as this would have additional

relevance for condensed matter. However, there are significant differences with the bosonic case. For instance, in the fermionic counterpart of the resource theory reported here, the free states, being both antisymmetric and particle-separable, would be just the single-particle and vacuum states.

It is hoped that the results presented here will stimulate further theoretical and experimental studies, across the communities of quantum information, quantum optics and condensed matter, in order to gain valuable insight into genuinely quantum properties of identical particles and their technological applications.

Chapter 6

Final remarks
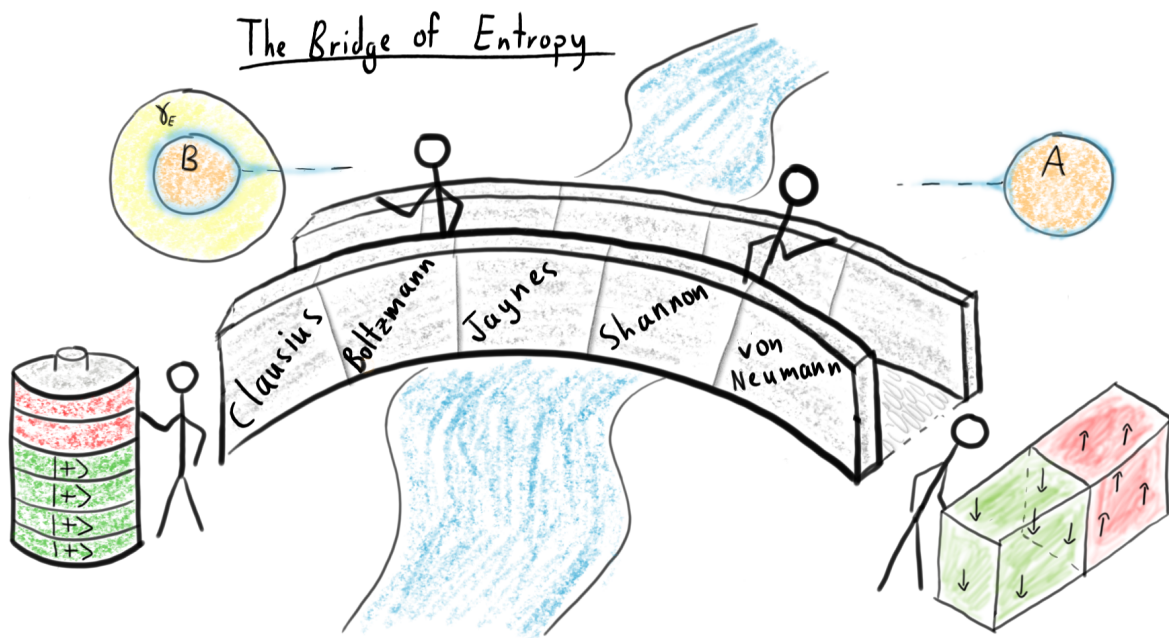
I started this thesis with a somewhat metaphorically strained bridge connecting together the two seemingly disparate topics of thermodynamics and quantum mechanics. In chapter 1 I presented how our calculation and understanding of the entropic function has provided the important conceptual and historical steps along the way that help us arrive at a consistent world view.

However, upon taking our steps along this bridge it quickly became apparent that although said bridge provided us with a way to cross over it did not tell us what to do when we got there. Hence why the title of this thesis was not simply "*The role of quantum information in thermodynamics*" (aside from the fact that this title was already taken [103]) but rather a *dialogue* between quantum information and thermodynamics, it

is very much a two way bridge.

In chapter 2, by elevating thermodynamics to a resource theory, with thermal non equilibrium being the resource in question, we are able to make equal handed statements about thermodynamic work and entanglement in the same equations. In chapter 3 *because* of results concerning the emergence of thermodynamic fluctuation relations from majorisation theory [8], we demonstrate an equivalent appearance of fluctuation relations governing the behaviour of quantum coherence.

In chapter 4, thanks to the conceptual notions of indistinguishability that arise from Gibbs' thermodynamic thought experiment, we are motivated to ask the same questions for quantum particles. Following on from this in chapter 5 we look more deeply at what this notion of indistinguishability means for the properties of an ensemble of identical quantum particles and the importance of the correcting $N!$ that was present in equivalent thermodynamic ensembles.

These works can only claim to be a result of the dialogue between thermodynamics and quantum not a supersedence of one physical theory by the other. In the future I hope that a better understanding of this complex relationship will not only allow us to state our field of *quantum thermodynamics* without ridicule, but will also give us greater insight into far larger questions: the emergence of the quantum from the classical, the way in which we understand and quantify information in a physical system, the incorporation or indeed separation of the observer in or from our theories. These are the very questions that arise during our metaphorical crossing.

I will finish with the following quote from (the supervisor of Charles Kittel) Frederic Keffer:

"The future belongs to those who can manipulate entropy; those who understand but energy will be only accountants." [108]

Bibliography

[1] ÅBERG, J. Catalytic coherence. *Physical Review Letters. 113* (Oct 2014), 150402.

[2] ABE, S., AND RAJAGOPAL, A. K. Validity of the second law in nonextensive quantum thermodynamics. *Physical Review Letters. 91* (Sep 2003), 120601.

[3] ABERG, J. Quantifying superposition. *arXiv preprint quant-ph/0612146* (2006).

[4] ÅBERG, J. Truly work-like work extraction via a single-shot analysis. *Nature Communications 4*, 1 (2013), 1–5.

[5] ADAMSON, R. B. A., TURNER, P. S., MITCHELL, M. W., AND STEINBERG, A. M. Detecting hidden differences via permutation symmetries. *Physical Review A 78*, 3 (2008), 033832.

[6] ADESSO, G., BROMLEY, T. R., AND CIANCIARUSO, M. Measures and applications of quantum correlations. *Journal of Physics A: Mathematical and Theoretical 49*, 47 (2016), 473001.

[7] ADESSO, G., BROMLEY, T. R., AND CIANCIARUSO, M. Measures and applications of quantum correlations. *Journal of Physics A 49*, 47 (2016), 473001.

[8] ALHAMBRA, Á. M., MASANES, L., OPPENHEIM, J., AND PERRY, C. Fluctuating work: From quantum thermodynamical identities to a second law equality. *Physical Review X 6*, 4 (2016), 041017.

[9] ALHAMBRA, Á. M., MASANES, L., OPPENHEIM, J., AND PERRY, C. Entanglement fluctuation theorems. *Physical Review A 100*, 1 (2019), 012317.

[10] ALLAHVERDYAN, A. E., AND NIEUWENHUIZEN, T. M. Explanation of the gibbs paradox within the framework of quantum thermodynamics. *Physical Review E 73*, 6 (2006), 066119.

[11] AN, S., ZHANG, J.-N., UM, M., LV, D., LU, Y., ZHANG, J., YIN, Z.-Q., QUAN, H., AND KIM, K. Experimental test of the quantum jarzynski equality with a trapped-ion system. *Nature Physics 11*, 2 (2015), 193–199.

[12] ANDERSON, M. H., ENSHER, J. R., MATTHEWS, M. R., WIEMAN, C. E., AND CORNELL, E. A. Observation of bose-einstein condensation in a dilute atomic vapor. *Science* (1995), 198–201.

[13] ASBÓTH, J. K., CALSAMIGLIA, J., AND RITSCH, H. Computable measure of nonclassicality for light. *Physical Review Letters. 94*, 17 (May 2005), 173602.

[14] AUDENAERT, K. M. A sharp continuity estimate for the von neumann entropy. *Journal of Physics A: Mathematical and Theoretical 40*, 28 (2007), 8127.

[15] AVERY, J. S. *Information theory and evolution.* World Scientific, 2012.

[16] BACH, A. *Indistinguishable Classical Particles.* Lecture Notes in Physics Monographs. Springer Berlin Heidelberg, 1996.

[17] BACH, A., AND LÜXMANN-ELLINGHAUS, U. The simplex structure of the classical states of the quantum harmonic oscillator. *Communications in Mathematical Physics 107*, 4 (dec 1986), 553–560.

[18] BACON, D., CHUANG, I. L., AND HARROW, A. W. Efficient quantum circuits for Schur and Clebsch-Gordan transforms. *Physical Review Letters 97*, 17 (2006), 170502.

[19] BALACHANDRAN, A., GOVINDARAJAN, T., DE QUEIROZ, A. R., AND REYES-LEGA, A. Entanglement and particle identity: a unifying approach. *Physical Review Letters 110*, 8 (2013), 080503.

[20] BARNUM, H., KNILL, E., ORTIZ, G., SOMMA, R., AND VIOLA, L. A subsystem-independent generalization of entanglement. *Physical Review Letters 92*, 10 (2004), 107902.

[21] BARNUM, H., ORTIZ, G., SOMMA, R., AND VIOLA, L. A generalization of entanglement to convex operational theories: entanglement relative to a subspace of observables. *International Journal of Theoretical Physics 44*, 12 (2005), 2127–2145.

[22] BARTLETT, S. D., RUDOLPH, T., AND SPEKKENS, R. W. Classical and quantum communication without a shared reference frame. *Physical Review Letters 91*, 2 (2003), 027901.

[23] BARTLETT, S. D., RUDOLPH, T., AND SPEKKENS, R. W. Reference frames, superselection rules, and quantum information. *Reviews of Modern Physics 79*, 2 (2007), 555.

[24] BAUMGRATZ, T., CRAMER, M., AND PLENIO, M. B. Quantifying coherence. *Physical Review Letters. 113* (Sep 2014), 140401.

[25] BENATTI, F., FLOREANINI, R., FRANCHINI, F., AND MARZOLINO, U. Remarks on entanglement and identical particles. *Open Systems & Information Dynamics 24*, 03 (2017), 1740004.

[26] BENATTI, F., FLOREANINI, R., FRANCHINI, F., AND MARZOLINO, U. Entanglement in indistinguishable particle systems. *Physics Reports* (2020).

[27] BENATTI, F., FLOREANINI, R., AND MARZOLINO, U. Entanglement and squeezing with identical particles: ultracold atom quantum metrology. *Journal of Physics B: Atomic, Molecular and Optical Physics 44*, 9 (2011), 091001.

[28] BENATTI, F., FLOREANINI, R., AND TITIMBO, K. Entanglement of identical particles. *Open Systems & Information Dynamics 21*, 01n02 (2014), 1440003.

[29] BENGTSSON, J., TENGSTRAND, M. N., WACKER, A., SAMUELSSON, P., UEDA, M., LINKE, H., AND REIMANN, S. Quantum Szilard engine with attractively interacting bosons. *Physical Review Letters 120*, 10 (2018), 100601.

[30] BENNETT, C. H. Notes on Landauer's principle, reversible computation, and Maxwell's Demon. *Studies In History and Philosophy of Science Part B: Studies In History and Philosophy of Modern Physics 34*, 3 (2003), 501–510.

[31] BENNETT, C. H., BRASSARD, G., CRÉPEAU, C., JOZSA, R., PERES, A., AND WOOTTERS, W. K. Teleporting an unknown quantum state via dual classical and einstein-podolsky-rosen channels. *Physical Review Letters. 70* (Mar 1993), 1895–1899.

[32] BENNETT, C. H., BRASSARD, G., POPESCU, S., SCHUMACHER, B., SMOLIN, J. A., AND WOOTTERS, W. K. Purification of noisy entanglement and faithful teleportation via noisy channels. *Physical Review Letters 76*, 5 (1996), 722.

[33] BENNETT, C. H., DIVINCENZO, D. P., SMOLIN, J. A., AND WOOTTERS, W. K. Mixed-state entanglement and quantum error correction. *Physical Review A 54*, 5 (1996), 3824.

[34] BENNETT, C. H., AND WIESNER, S. J. Communication via one-and two-particle operators on einstein-podolsky-rosen states. *Physical Review Letters 69*, 20 (1992), 2881.

[35] BERA, M. N., RIERA, A., LEWENSTEIN, M., AND WINTER, A. Generalized laws of thermodynamics in the presence of correlations. *Nature Communications 8* (2017), 2180.

[36] BERGMANN, P. G., AND LEBOWITZ, J. L. New approach to nonequilibrium processes. *Physical Review 99*, 2 (1955), 578.

[37] BHATIA, R. *Matrix Analysis.* Springer, 1996.

[38] BIEDENHARN, L. C., AND LOUCK, J. D. *Angular Momentum in Quantum Physics.* Addison-Wesley, Reading, MA, 1981.

[39] BINDER, F., CORREA, L. A., GOGOLIN, C., ANDERS, J., AND ADESSO, G., Eds. *Thermodynamics in the quantum regime.* Fundamental Theories of Physics. Springer, 2019.

[40] BOLTZMANN, L. Über die Beziehung zwischen dem zweiten Hauptsatze der mechanischen Wärmetheorie und der Wahrscheinlichkeitsrechnung respektive den Sätzen über das Wärmegleichgewicht. *Wiener Berichte 76* (1877), 373–435.

[41] BRANDÃO, F. G. S. L., HORODECKI, M., OPPENHEIM, J., RENES, J. M., AND SPEKKENS, R. W. Resource theory of quantum states out of thermal equilibrium. *Physical Review Letters. 111* (Dec 2013), 250404.

[42] BRANDAO, F. G. Quantifying entanglement with witness operators. *Physical Review A 72*, 2 (2005), 022310.

[43] BRANDAO, F. G., HORODECKI, M., OPPENHEIM, J., RENES, J. M., AND SPEKKENS, R. W. Resource theory of quantum states out of thermal equilibrium. *Physical Review Letters 111*, 25 (2013), 250404.

[44] BRANDÃO, F. G. S. L., HORODECKI, M., NG, N. H. Y., OPPENHEIM, J., AND WEHNER, S. The second laws of quantum thermodynamics. *Proceedings of the National Academy of Sciences of the United States of America 112*, 11 (2015), 3275–3279.

[45] BRANDNER, K., AND SAITO, K. Thermodynamic geometry of microscopic heat engines. *Physical review letters 124*, 4 (2020), 040602.

[46] BRAUN, D., ADESSO, G., BENATTI, F., FLOREANINI, R., MARZOLINO, U., MITCHELL, M. W., AND PIRANDOLA, S. Quantum-enhanced measurements without entanglement. *Reviews of Modern Physics 90*, 3 (sep 2018), 035006.

[47] BRUSCHI, D. E., DRAGAN, A., FUENTES, I., AND LOUKO, J. Particle and antiparticle bosonic entanglement in noninertial frames. *Physical Review D 86*, 2 (2012), 025026.

[48] BRUSCHI, D. E., MORRIS, B., AND FUENTES, I. Thermodynamics of relativistic quantum fields confined in cavities. *Physics Letters A* (2020), 126601.

[49] Bu, K., Singh, U., and Wu, J. Catalytic coherence transformations. *Physical Review A 93* (Apr 2016), 042326.

[50] Carnot, S. Réflexions sur la puissance motrice du feu et sur les machines propres à développer cette puissance. Annales scientifiques de l'Ecole normale.

[51] Castin, Y., and Dum, R. Bose-einstein condensates in time dependent traps. *Physical Review Letters 77* (Dec 1996), 5315–5319.

[52] Cavalcanti, D., Malard, L., Matinaga, F., Cunha, M. T., and Santos, M. F. Useful entanglement from the pauli principle. *Physical Review B 76*, 11 (2007), 113304.

[53] Caves, C. M., Fuchs, C. A., and Schack, R. Unknown quantum states: the quantum de finetti representation. *Journal of Mathematical Physics 43*, 9 (2002), 4537–4559.

[54] Chitambar, E., and Gour, G. Comparison of incoherent operations and measures of coherence. *Physical Review A 94* (Nov 2016), 052336.

[55] Chitambar, E., and Gour, G. Critical examination of incoherent operations and a physically consistent resource theory of quantum coherence. *Physical Review Letters. 117* (Jul 2016), 030401.

[56] Chitambar, E., and Gour, G. Erratum: Comparison of incoherent operations and measures of coherence [physical review a 94, 052336 (2016)]. *Physical Review A 95* (Jan 2017), 019902.

[57] Chitambar, E., and Gour, G. Quantum resource theories. *Reviews of Modern Physics 91*, 2 (2019), 025001.

[58] Chitambar, E., Streltsov, A., Rana, S., Bera, M., Adesso, G., and Lewenstein, M. Assisted distillation of quantum coherence. *Physical Review Letters 116*, 7 (2016), 070402.

[59] Clausius, R. "U about a modified ä form of the second law of the mechanical w "a rmetheorie. *Annals of Physics 169*, 12 (1854), 481–506.

[60] Collin, D., Ritort, F., Jarzynski, C., Smith, S. B., Tinoco Jr, I., and Bustamante, C. Verification of the crooks fluctuation theorem and recovery of rna folding free energies. *Nature 437*, 7056 (2005), 231.

[61] Compagno, G., Castellini, A., and Lo Franco, R. Dealing with indistinguishable particles and their entanglement. *Philosophical Transactions of the Royal Society A 376*, 2123 (2018), 20170317.

[62] CORREA, L. A., XU, B., MORRIS, B., AND ADESSO, G. Pushing the limits of the reaction-coordinate mapping. *The Journal of chemical physics 151*, 9 (2019), 094107.

[63] COTTET, N., JEZOUIN, S., BRETHEAU, L., CAMPAGNE-IBARCQ, P., FICHEUX, Q., ANDERS, J., AUFFÈVES, A., AZOUIT, R., ROUCHON, P., AND HUARD, B. Observing a quantum maxwell demon at work. *Proceedings of the National Academy of Sciences 114*, 29 (2017), 7561–7564.

[64] CROOKS, G. E. Entropy production fluctuation theorem and the nonequilibrium work relation for free energy differences. *Physical Review E 60*, 3 (1999), 2721.

[65] DAHLSTEN, O. C. O., RENNER, R., RIEPER, E., AND VEDRAL, V. Inadequacy of von Neumann entropy for characterizing extractable work. *New Journal of Physics 13*, 5 (may 2011), 053015.

[66] DALTON, B., HEANEY, L., GOOLD, J., GARRAWAY, B., AND BUSCH, T. New spin squeezing and other entanglement tests for two mode systems of identical bosons. *New Journal of Physics 16*, 1 (2014), 013026.

[67] DALTON, B. J., GOOLD, J., GARRAWAY, B. M., AND REID, M. D. Quantum entanglement for systems of identical bosons: I. General features. *Physica Scripta 92*, 2 (feb 2017), 023004.

[68] DARRIGOL, O. The Gibbs paradox: Early history and solutions. *Entropy 20*, 6 (2018), 443.

[69] DEFFNER, S., AND LUTZ, E. Generalized clausius inequality for nonequilibrium quantum processes. *Physical Review Letters 105*, 17 (2010), 170402.

[70] DEVETAK, I. Distillation of local purity from quantum states. *Physical Review A 71*, 6 (2005), 062303.

[71] DEVETAK, I., AND BERGER, T. Low-entanglement remote state preparation. *Physical Review Letters 87*, 19 (2001), 197901.

[72] DEVETAK, I., AND WINTER, A. Distilling common randomness from bipartite quantum states. *IEEE Transactions on Information Theory 50*, 12 (2004), 3183–3196.

[73] DIEKS, D. The Gibbs Paradox and Particle Individuality. *Entropy 20*, 6 (jun 2018), 466.

[74] DIRAC, P. A. M. *The principles of quantum mechanics.* No. 27. Oxford university press, 1981.

[75] DiVincenzo, D. P., Fuchs, C. A., Mabuchi, H., Smolin, J. A., Thapliyal, A., and Uhlmann, A. Entanglement of assistance. In *NASA International Conference on Quantum Computing and Quantum Communications* (1998), Springer, pp. 247–257.

[76] Du, S., Bai, Z., and Guo, Y. Conditions for coherence transformations under incoherent operations. *Physical Review A 91* (May 2015), 052120.

[77] Du, S., Bai, Z., and Guo, Y. Erratum: Conditions for coherence transformations under incoherent operations [physical review a 91, 052120 (2015)]. *Physical Review A 95* (Feb 2017), 029901.

[78] Eckert, K., Schliemann, J., Bruss, D., and Lewenstein, M. Quantum correlations in systems of indistinguishable particles. *Annals of Physics 299*, 1 (2002), 88–127.

[79] Ekert, A. K. Quantum cryptography based on bells theorem. *Physical Review Letters 67*, 6 (1991), 661.

[80] Elouard, C., Herrera-Martí, D. A., Clusel, M., and Auffèves, A. The role of quantum measurement in stochastic thermodynamics. *NPJ Quantum Information 3*, 1 (2017), 9.

[81] Fadel, M., Zibold, T., Décamps, B., and Treutlein, P. Spatial entanglement patterns and einstein-podolsky-rosen steering in bose-einstein condensates. *Science 360*, 6387 (2018), 409–413.

[82] Faist, P., Oppenheim, J., and Renner, R. Gibbs-preserving maps outperform thermal operations in the quantum regime. *New Journal of Physics 17*, 4 (2015), 043003.

[83] Fannes, M. A continuity property of the entropy density for spin lattice systems. *Communications in Mathematical Physics 31*, 4 (1973), 291–294.

[84] Feynman, R., Leighton, R., and Sands, M. In *The Feynman Lectures on Physics, Vol. III: The New Millennium Edition: Quantum Mechanics*, The Feynman Lectures on Physics. Basic Books, 2011, ch. 4.

[85] Flajolet, P., and Sedgewick, R. *Analytic Combinatorics*. Cambridge University Press, Cambridge, 2009.

[86] Francica, G., Goold, J., Plastina, F., and Paternostro, M. Daemonic ergotropy: enhanced work extraction from quantum correlations. *NPJ Quantum Information 3*, 1 (2017), 12.

[87] FRÖWIS, F., SEKATSKI, P., DÜR, W., GISIN, N., AND SANGOUARD, N. Macroscopic quantum states: Measures, fragility, and implementations. *Reviews of Modern Physics 90*, 2 (may 2018), 025004.

[88] FUJITA, S. On the indistinguishability of classical particles. *Foundations of Physics 21*, 4 (1991), 439–457.

[89] GALLEGO, R., EISERT, J., AND WILMING, H. Thermodynamic work from operational principles. *New Journal of Physics 18*, 10 (2016), 103017.

[90] GARBE, L., FELICETTI, S., MILMAN, P., COUDREAU, T., AND KELLER, A. Metrological advantage at finite temperature for gaussian phase estimation. *Physical Review A 99*, 4 (2019), 043815.

[91] GEHRKE, C., SPERLING, J., AND VOGEL, W. Quantification of nonclassicality. *Physical Review A 86*, 5 (nov 2012), 052118.

[92] GERRY, C., AND KNIGHT, P. *Introductory Quantum Optics*. Cambridge University Press, Cambridge, 2004.

[93] GESSNER, M., PEZZÈ, L., AND SMERZI, A. Resolution-enhanced entanglement detection. *Physical Review A 95*, 3 (mar 2017), 032326.

[94] GHIRARDI, G., AND MARINATTO, L. General criterion for the entanglement of two indistinguishable particles. *Physical Review A 70*, 1 (2004), 012109.

[95] GHIRARDI, G., MARINATTO, L., AND WEBER, T. Entanglement and properties of composite quantum systems: a conceptual and mathematical analysis. *Journal of Statistical Physics 108*, 1-2 (2002), 49–122.

[96] GHIRARDI, G. C., RIMINI, A., WEBER, T., AND OMERO, C. Some simple remarks about quantum nonseparability for systems composed of identical constituents. *Il Nuovo Cimento B (1971-1996) 39*, 1 (1977), 130–134.

[97] GIBBS, J. W. On the equilibrium of heterogeneous substances. *Transactions of the Connecticut Academy of Arts and Sciences* (1879).

[98] GIOVANNETTI, V., LLOYD, S., AND MACCONE, L. Quantum metrology. *Physical Review Letters 96*, 1 (2006), 010401.

[99] GIOVANNETTI, V., MANCINI, S., VITALI, D., AND TOMBESI, P. Characterizing the entanglement of bipartite quantum systems. *Physical Review A 67*, 2 (2003), 022320.

[100] GIRAUD, O., BRAUN, P., AND BRAUN, D. Quantifying quantumness and the quest for Queens of Quantum. *New Journal of Physics 12*, 6 (jun 2010), 063005.

[101] GLAUBER, R. J. The Quantum Theory of Optical Coherence. *Physical REview 130*, 6 (jun 1963), 2529.

[102] GOODMAN, R., AND WALLACH, N. R. *Symmetry, Representations, and Invariants*, vol. 255 of *Graduate Texts in Mathematics*. Springer New York, New York, NY, 2009.

[103] GOOLD, J., HUBER, M., RIERA, A., DEL RIO, L., AND SKRZYPCZYK, P. The role of quantum information in thermodynamics — a topical review. *Journal of Physics A: Mathematical and Theoretical 49*, 14 (2016), 143001.

[104] GOUR, G., MARVIAN, I., AND SPEKKENS, R. W. Measuring the quality of a quantum reference frame: The relative entropy of frameness. *Physical Review A 80* (Jul 2009), 012307.

[105] GOUR, G., MÜLLER, M. P., NARASIMHACHAR, V., SPEKKENS, R. W., AND HALPERN, N. Y. The resource theory of informational nonequilibrium in thermodynamics. *Physics Reports 583* (2015), 1–58.

[106] GROSS, C., ZIBOLD, T., NICKLAS, E., ESTEVE, J., AND OBERTHALER, M. K. Nonlinear atom interferometer surpasses classical precision limit. *Nature 464*, 7292 (2010), 1165.

[107] GUARNIERI, G., NG, N., MODI, K., EISERT, J., PATERNOSTRO, M., AND GOOLD, J. Quantum work statistics and resource theories: bridging the gap through rényi divergences. *Physical Review E 99*, 5 (2019), 050101.

[108] HADDAD, W. M., CHELLABOINA, V., AND NERSESOV, S. G. *Thermodynamics: A dynamical systems approach*, vol. 23. Princeton University Press, 2009.

[109] HALPERN, N. Y., GARNER, A. J. P., DAHLSTEN, O. C. O., AND VEDRAL, V. What's the worst that could happen? one-shot dissipated work from rényi divergences. *arXiv preprint arXiv:1505.06217* (2015).

[110] HAMERMESH, M. *Group Theory and Its Application to Physical Problems*. Addison Wesley Series in Physics. Dover Publications, 1989.

[111] HARREMOES, P., AND RUZANKIN, P. Rate of Convergence to Poisson Law in Terms of Information Divergence. *IEEE Transactions on Information Theory 50*, 9 (sep 2004), 2145–2149.

[112] HARROW, A. W. Applications of coherent classical communication and the schur transform to quantum information theory. *arXiv preprint quant-ph/0512255* (2005).

[113] HARROW, A. W. The church of the symmetric subspace. *arXiv preprint arXiv:1308.6595* (2013).

[114] HASTINGS, M. B. Superadditivity of communication capacity using entangled inputs. *Nature Physics 5*, 4 (2009), 255–257.

[115] HAYASHI, M. *Quantum Information Theory: Mathematical Foundation.* Graduate Texts in Physics. Springer Berlin Heidelberg, 2016.

[116] HAYDEN, P. M., HORODECKI, M., AND TERHAL, B. M. The asymptotic entanglement cost of preparing a quantum state. *Journal of Physics A: Mathematical and General 34*, 35 (2001), 6891.

[117] HENDERSON, L., AND VEDRAL, V. Classical, quantum and total correlations. *Journal of physics A: mathematical and general 34*, 35 (2001), 6899.

[118] HERBUT, F. A quantum measure of coherence and incompatibility. *Journal of Physics A 38*, 13 (2005), 2959.

[119] HERDMAN, C., ROY, P.-N., MELKO, R., AND DEL MAESTRO, A. Entanglement area law in superfluid 4 he. *Nature Physics 13*, 6 (2017), 556.

[120] HOLIAN, B. L., HOOVER, W. G., AND POSCH, H. A. Resolution of Loschmidts paradox: The origin of irreversible behavior in reversible atomistic dynamics. *Physical Review Letters 59*, 1 (1987), 10.

[121] HOLMES, Z., ANDERS, J., AND MINTERT, F. Enhanced energy transfer to an optomechanical piston from indistinguishable photons. *Physical Review Letters 124*, 21 (2020), 210601.

[122] HOLMES, Z., ANDERS, J., AND MINTERT, F. Gibbs mixing of partially distinguishable photons with a polarising beamsplitter membrane. *arXiv preprint arXiv:2006.00613* (2020).

[123] HONG, C. K., OU, Z. Y., AND MANDEL, L. Measurement of subpicosecond time intervals between two photons by interference. *Physical Review Letters 59*, 18 (nov 1987), 2044–2046.

[124] HORODECKI, M., HORODECKI, K., HORODECKI, P., HORODECKI, R., OPPENHEIM, J., SEN, A., SEN, U., ET AL. Local information as a resource in distributed quantum systems. *Physical Review Letters 90*, 10 (2003), 100402.

[125] HORODECKI, M., AND HORODECKI, P. Reduction criterion of separability and limits for a class of distillation protocols. *Physical Review A 59*, 6 (1999), 4206.

[126] HORODECKI, M., HORODECKI, P., AND HORODECKI, R. Mixed-state entanglement and distillation: Is there a bound entanglement in nature? *Physical Review Letters 80*, 24 (1998), 5239.

[127] HORODECKI, M., AND OPPENHEIM, J. Fundamental limitations for quantum and nanoscale thermodynamics. *Nature Communications 4* (2013), 2059.

[128] HORODECKI, R., HORODECKI, P., HORODECKI, M., AND HORODECKI, K. Quantum entanglement. *Reviews of Modern Physics. 81* (Jun 2009), 865–942.

[129] HYLLUS, P., LASKOWSKI, W., KRISCHEK, R., SCHWEMMER, C., WIECZOREK, W., WEINFURTER, H., PEZZÉ, L., AND SMERZI, A. Fisher information and multiparticle entanglement. *Physical Review A 85*, 2 (Feb 2012), 022321.

[130] HYLLUS, P., PEZZÉ, L., SMERZI, A., AND TÓTH, G. Entanglement and extreme spin squeezing for a fluctuating number of indistinguishable particles. *Physical Review A 86*, 1 (2012), 012337.

[131] ICHIKAWA, T., SASAKI, T., TSUTSUI, I., AND YONEZAWA, N. Exchange symmetry and multipartite entanglement. *Physical Review A 78*, 5 (2008), 052105.

[132] JARZYNSKI, C. Nonequilibrium equality for free energy differences. *Physical Review Letters. 78*, 14 (1997), 2690.

[133] JAYNES, E. T. Information theory and statistical mechanics. *Physical review 106*, 4 (1957), 620.

[134] JAYNES, E. T. The gibbs paradox. In *Maximum entropy and bayesian methods*. Springer, 1992, pp. 1–21.

[135] JONES, S. J., WISEMAN, H. M., BARTLETT, S., VACCARO, J. A., AND POPE, D. T. Entanglement and symmetry: A case study in superselection rules, reference frames, and beyond. *Physical Review A 74*, 6 (2006), 062313.

[136] KAUFMAN, A. M., TAI, M. E., LUKIN, A., RISPOLI, M., SCHITTKO, R., PREISS, P. M., AND GREINER, M. Quantum thermalization through entanglement in an isolated many-body system. *Science 353*, 6301 (2016), 794–800.

[137] KEYL, M., AND WERNER, R. F. Optimal cloning of pure states, testing single clones. *Journal of Mathematical Physics 40*, 7 (1999), 3283.

[138] KILLORAN, N., CRAMER, M., AND PLENIO, M. B. Extracting entanglement from identical particles. *Physical Review Letters 112*, 15 (2014), 150501.

[139] KILLORAN, N., STEINHOFF, F. E. S., AND PLENIO, M. B. Converting nonclassicality into entanglement. *Physical Review Letters. 116*, 8 (Feb 2016), 080402.

[140] KIM, M. S., SON, W., BUŽEK, V., AND KNIGHT, P. L. Entanglement by a beam splitter: Nonclassicality as a prerequisite for entanglement. *Physical Review A 65*, 3 (Feb 2002), 032323.

[141] KIM, S. W., SAGAWA, T., DE LIBERATO, S., AND UEDA, M. Quantum szilard engine. *Physical Review Letters 106*, 7 (2011), 070401.

[142] KING, C. The capacity of the quantum depolarizing channel. *IEEE Transactions on Information Theory 49*, 1 (2003), 221–229.

[143] KOASHI, M., AND WINTER, A. Monogamy of quantum entanglement and other correlations. *Physical Review A 69*, 2 (2004), 022309.

[144] KORZEKWA, K., LOSTAGLIO, M., OPPENHEIM, J., AND JENNINGS, D. The extraction of work from quantum coherence. *New Journal of Physics 18*, 2 (2016), 023045.

[145] KOSLOFF, R. Quantum thermodynamics: A dynamical viewpoint. *Entropy 15*, 6 (2013), 2100.

[146] KUNKEL, P., PRÜFER, M., STROBEL, H., LINNEMANN, D., FRÖLIAN, A., GASENZER, T., GÄRTTNER, M., AND OBERTHALER, M. K. Spatially distributed multipartite entanglement enables epr steering of atomic clouds. *Science 360*, 6387 (2018), 413–416.

[147] KWON, H., TAN, K. C., VOLKOFF, T., AND JEONG, H. Nonclassicality as a Quantifiable Resource for Quantum Metrology. *Physical Review Letters 122*, 4 (feb 2019), 040503.

[148] LAMI, L. Completing the grand tour of asymptotic quantum coherence manipulation. *IEEE Transactions on Information Theory 66*, 4 (2019), 2165–2183.

[149] LAMI, L., REGULA, B., AND ADESSO, G. Generic bound coherence under strictly incoherent operations. *Physical Review Letters 122*, 15 (2019), 150402.

[150] LANDAU, L., AND LIFSHITZ, E. *Statistical Physics: Volume 5*. No. v. 5. Elsevier Science, 2013.

[151] LANDAUER, R. Irreversibility and heat generation in the computing process. *IBM journal of research and development 5*, 3 (1961), 183–191.

[152] LANDI, G. T., GUARNIERI, G., MORRIS, B., GOOLD, J., AND ADESSO, G. A resource theory of maxwell's demons. *arXiv preprint arXiv:1903.12629* (2019).

[153] LANGE, K., PEISE, J., LÜCKE, B., KRUSE, I., VITAGLIANO, G., APELLANIZ, I., KLEINMANN, M., TÓTH, G., AND KLEMPT, C. Entanglement between two spatially separated atomic modes. *Science 360*, 6387 (2018), 416–418.

[154] LEVITIN, L. B. Gibbs paradox and equivalence relation between quantum information and work. In *Workshop on Physics and Computation* (1992), IEEE, pp. 223–226.

[155] LIUZZO-SCORPO, P., CORREA, L. A., SCHMIDT, R., AND ADESSO, G. Thermodynamics of quantum feedback cooling. *Entropy 18*, 2 (2016), 48.

[156] LLOYD, S. A potentially realizable quantum computer. *Science 261*, 5128 (1993), 1569–1571.

[157] LLOYD, S., AND BRAUNSTEIN, S. L. Quantum computation over continuous variables. In *Quantum Information with Continuous Variables.* Springer, 1999, pp. 9–17.

[158] LO FRANCO, R., AND COMPAGNO, G. Quantum entanglement of identical particles by standard information-theoretic notions. *Scientific Reports 6* (2016), 20603.

[159] LO FRANCO, R., AND COMPAGNO, G. Indistinguishability of elementary systems as a resource for quantum information processing. *Physical Review Letters 120*, 24 (2018), 240403.

[160] LOSTAGLIO, M., JENNINGS, D., AND RUDOLPH, T. Description of quantum coherence in thermodynamic processes requires constraints beyond free energy. *Nature Communications 6* (mar 2015), 6383.

[161] MA, J., YADIN, B., GIROLAMI, D., VEDRAL, V., AND GU, M. Converting Coherence to Quantum Correlations. *Physical Review Letters 116*, 16 (2016), 160407.

[162] MANDL, F. *Statistical Physics.* Manchester Physics Series. Wiley, 2013.

[163] MANZANO, G., PLASTINA, F., AND ZAMBRINI, R. Optimal work extraction and thermodynamics of quantum measurements and correlations. *Physical Review Letters. 121* (Sep 2018), 120602.

[164] MARTÍNEZ, I. A., ROLDÁN, É., DINIS, L., PETROV, D., AND RICA, R. A. Adiabatic processes realized with a trapped brownian particle. *Physical Review Letters 114*, 12 (2015), 120601.

[165] MARVIAN, I., AND SPEKKENS, R. W. How to quantify coherence: Distinguishing speakable and unspeakable notions. *Physical Review A 94* (Nov 2016), 052324.

[166] MASANES, L., AND OPPENHEIM, J. A general derivation and quantification of the third law of thermodynamics. *Nature Communications 8* (2017), 14538.

[167] MOHAMMADY, M. H., AND ANDERS, J. A quantum szilard engine without heat from a thermal reservoir. *New Journal of Physics 19*, 11 (2017), 113026.

[168] MORRIS, B., AND ADESSO, G. Quantum coherence fluctuation relations. *Journal of Physics A: Mathematical and Theoretical 51*, 41 (2018), 414007.

[169] MORRIS, B., LAMI, L., AND ADESSO, G. Assisted work distillation. *Physical Review Letters 122*, 13 (2019), 130601.

[170] MORRIS, B., YADIN, B., FADEL, M., ZIBOLD, T., TREUTLEIN, P., AND ADESSO, G. Entanglement between identical particles is a useful and consistent resource. *Physical Review X 10* (Oct 2020), 041012.

[171] MÜLLER, M. P. Correlating thermal machines and the second law at the nanoscale. *Physical Review X 8*, 4 (2018), 041051.

[172] MYERS, N. M., AND DEFFNER, S. Bosons outperform fermions: The thermodynamic advantage of symmetry. *Physical Review E 101*, 1 (jan 2020), 012110.

[173] NAPOLI, C., BROMLEY, T. R., CIANCIARUSO, M., PIANI, M., JOHNSTON, N., AND ADESSO, G. Robustness of coherence: An operational and observable measure of quantum coherence. *Physical Review Letters. 116* (Apr 2016), 150502.

[174] NARASIMHACHAR, V., ASSAD, S., BINDER, F. C., THOMPSON, J., YADIN, B., AND GU, M. Thermodynamic resources in continuous-variable quantum systems. *arXiv:1909.07364* (2019).

[175] NARASIMHACHAR, V., AND GOUR, G. Resource theory under conditioned thermal operations. *Physical Review A 95*, 1 (2017), 012313.

[176] NARASIMHACHAR, V., THOMPSON, J., MA, J., GOUR, G., AND GU, M. Quantifying memory capacity as a quantum thermodynamic resource. *Physical Review Letters 122*, 6 (2019), 060601.

[177] NEUMANN, J. *Mathematische grundlagen der quantenmechanik*. Verlag von Julius Springer Berlin, 1932.

[178] NIEDENZU, W., HUBER, M., AND BOUKOBZA, E. Concepts of work in autonomous quantum heat engines. *Quantum 3* (2019), 195.

[179] NIELSEN, M. A. Conditions for a class of entanglement transformations. *Physical Review Letters. 83*, 2 (1999), 436.

[180] NIELSEN, M. A., AND CHUANG, I. Quantum computation and quantum information, 2002.

[181] NIELSEN, M. A., AND VIDAL, G. Majorization and the Interconversion of Bipartite States. *Quantum Information and Computation 1* (2001), 76–93.

[182] OCKELOEN, C. F., SCHMIED, R., RIEDEL, M. F., AND TREUTLEIN, P. Quantum metrology with a scanning probe atom interferometer. *Physical review letters 111*, 14 (2013), 143001.

[183] OLLIVIER, H., AND ZUREK, W. H. Quantum discord: a measure of the quantumness of correlations. *Physical Review Letters 88*, 1 (2001), 017901.

[184] OPPENHEIM, J., HORODECKI, M., HORODECKI, P., AND HORODECKI, R. Thermodynamical approach to quantifying quantum correlations. *Physical Review Letters 89*, 18 (2002), 180402.

[185] PARIS, M. G. A. Quantum estimation for quantum technology. *International Journal of Quantum Information. 7*, supp01 (jan 2009), 125.

[186] PARK, D. K., RODRIGUEZ-BRIONES, N. A., FENG, G., RAHIMI, R., BAUGH, J., AND LAFLAMME, R. Heat bath algorithmic cooling with spins: review and prospects. *Electron Spin Resonance (ESR) Based Quantum Computing* (2016), 227–255.

[187] PARK, J. J., KIM, K.-H., SAGAWA, T., AND KIM, S. W. Heat engine driven by purely quantum information. *Physical Review Letters 111*, 23 (2013), 230402.

[188] PAŠKAUSKAS, R., AND YOU, L. Quantum correlations in two-boson wave functions. *Physical Review A 64*, 4 (2001), 042310.

[189] PETROV, V. V. *Sums of Independent Random Variables.* Springer Berlin Heidelberg, Berlin, Heidelberg, 1975.

[190] PETZ, D. 'E. N. Entropy, von neumann and the von neumann entropy. In *John von Neumann and the foundations of quantum physics.* Springer, 2001, pp. 83–96.

[191] PEZZÈ, L., SMERZI, A., OBERTHALER, M. K., SCHMIED, R., AND TREUTLEIN, P. Quantum metrology with nonclassical states of atomic ensembles. *Reviews of Modern Physics 90*, 3 (2018), 035005.

[192] PIANI, M., GHARIBIAN, S., ADESSO, G., CALSAMIGLIA, J., HORODECKI, P., AND WINTER, A. All nonclassical correlations can be activated into distillable entanglement. *Physical Review Letters. 106*, 22 (Jun 2011), 220403.

[193] PLESCH, M., DAHLSTEN, O., GOOLD, J., AND VEDRAL, V. Maxwell's daemon: Information versus particle statistics. *Scientific reports 4* (2014), 6995.

[194] POPOVICIU, T. Sur les quations algbriques ayant toutes leurs racines relles. *Mathematica 9* (1935), 129 to 145.

[195] RECK, M., ZEILINGER, A., BERNSTEIN, H. J., AND BERTANI, P. Experimental realization of any discrete unitary operator. *Physical Review Letters 73*, 1 (jul 1994), 58–61.

[196] REEB, D., AND WOLF, M. M. An improved landauer principle with finite-size corrections. *New Journal of Physics 16*, 10 (2014), 103011.

[197] REGULA, B., FANG, K., WANG, X., AND ADESSO, G. One-shot coherence distillation. *Physical Review Letters 121*, 1 (2018), 010401.

[198] REGULA, B., LAMI, L., AND STRELTSOV, A. Nonasymptotic assisted distillation of quantum coherence. *Physical Review A 98* (2018), 052329.

[199] REUSCH, A., SPERLING, J., AND VOGEL, W. Entanglement witnesses for indistinguishable particles. *Physical Review A 91*, 4 (2015), 042324.

[200] RIEDEL, M. F., BÖHI, P., LI, Y., HÄNSCH, T. W., SINATRA, A., AND TREUTLEIN, P. Atom-chip-based generation of entanglement for quantum metrology. *Nature 464*, 7292 (2010), 1170.

[201] SAGAWA, T., AND UEDA, M. Second law of thermodynamics with discrete quantum feedback control. *Physical Review Letters 100*, 8 (2008), 080403.

[202] SASAKI, T., ICHIKAWA, T., AND TSUTSUI, I. Entanglement of indistinguishable particles. *Physical Review A 83*, 1 (2011), 012113.

[203] SAUNDERS, S. The Gibbs Paradox. *Entropy 20*, 8 (jul 2018), 552.

[204] SCHUCH, N., VERSTRAETE, F., AND CIRAC, J. I. Nonlocal resources in the presence of superselection rules. *Physical Review Letters 92*, 8 (2004), 087904.

[205] SCHUCH, N., VERSTRAETE, F., AND CIRAC, J. I. Quantum entanglement theory in the presence of superselection rules. *Physical Review A 70*, 4 (2004), 042310.

[206] SEARS, F. W. Principles of physics i: Mechanics heat and sound. Tech. rep., 1947.

[207] SEKIMOTO, K. *Stochastic energetics*, vol. 799. Springer, 2010.

[208] SERAFINI, A., LOSTAGLIO, M., LONGDEN, S., SHACKERLEY-BENNETT, U., HSIEH, C.-Y., AND ADESSO, G. Gaussian Thermal Operations and The Limits of Algorithmic Cooling. *Physical Review Letters 124*, 1 (jan 2020), 010602.

[209] SHANNON, C. E. A mathematical theory of communication. *ACM SIGMOBILE Mobile Computing and Communications Review 5*, 1 (2001), 3–55.

[210] SHI, Y. Quantum entanglement of identical particles. *Physical Review A 67*, 2 (2003), 024301.

[211] SHOR, P. W., AND PRESKILL, J. Simple proof of security of the bb84 quantum key distribution protocol. *Physical Review Letters 85*, 2 (2000), 441.

[212] SIMON, C. Natural entanglement in bose-einstein condensates. *Physical Review A 66*, 5 (2002), 052323.

[213] SØRENSEN, A., DUAN, L.-M., CIRAC, J., AND ZOLLER, P. Many-particle entanglement with bose–einstein condensates. *Nature 409*, 6816 (2001), 63.

[214] SPERLING, J., AND VOGEL, W. The Schmidt number as a universal entanglement measure. *Physica Scripta 83*, 4 (apr 2011), 045002.

[215] STANISIC, S., AND TURNER, P. S. Discriminating distinguishability. *Physical Review A 98*, 4 (oct 2018), 043839.

[216] STRELTSOV, A., ADESSO, G., AND PLENIO, M. B. Colloquium: quantum coherence as a resource. *Reviews of Modern Physics 89*, 4 (2017), 041003.

[217] STRELTSOV, A., RANA, S., BERA, M. N., AND LEWENSTEIN, M. Towards resource theory of coherence in distributed scenarios. *Physical Review X 7*, 1 (2017), 011024.

[218] STRELTSOV, A., SINGH, U., DHAR, H. S., BERA, M. N., AND ADESSO, G. Measuring quantum coherence with entanglement. *Physical Review Letters. 115* (Jul 2015), 020403.

[219] STRELTSOV, A., SINGH, U., DHAR, H. S., BERA, M. N., AND ADESSO, G. Measuring quantum coherence with entanglement. *Physical Review Letters 115*, 2 (2015), 020403.

[220] STROBEL, H., MUESSEL, W., LINNEMANN, D., ZIBOLD, T., HUME, D. B., PEZZE, L., SMERZI, A., AND OBERTHALER, M. K. Fisher information and entanglement of non-Gaussian spin states. *Science 345*, 6195 (jul 2014), 424–427.

[221] SUDARSHAN, E. C. G. Equivalence of semiclassical and quantum mechanical descriptions of statistical light beams. *Physical Review Letters. 10*, 7 (Apr 1963), 277–279.

[222] SZILARD, L. Über die entropieverminderung in einem thermodynamischen system bei eingriffen intelligenter wesen. *Zeitschrift für Physik 53*, 11-12 (1929), 840–856.

[223] TAN, K. C., VOLKOFF, T., KWON, H., AND JEONG, H. Quantifying the coherence between coherent states. *Physical Review Letters 119*, 19 (2017), 190405.

[224] TICHY, M. C., DE MELO, F., KUŚ, M., MINTERT, F., AND BUCHLEITNER, A. Entanglement of identical particles and the detection process. *Fortschritte der Physik 61*, 2-3 (2013), 225–237.

[225] TICHY, M. C., MINTERT, F., AND BUCHLEITNER, A. Essential entanglement for atomic and molecular physics. *Journal of Physics B: Atomic, Molecular and Optical Physics 44*, 19 (2011), 192001.

[226] TÓTH, G. Multipartite entanglement and high-precision metrology. *Physical Review A 85*, 2 (Feb 2012), 022322.

[227] TOYABE, S., AND SANO, M. Nonequilibrium fluctuations in biological strands, machines, and cells. *Journal of the Physical Society of Japan 84*, 10 (2015), 102001.

[228] TUCKERMAN, M. Statistical mechanics: theory and molecular simulation. Oxford university press, 2010, ch. 11.

[229] UZDIN, R., LEVY, A., AND KOSLOFF, R. Equivalence of quantum heat machines, and quantum-thermodynamic signatures. *Physical Review X 5* (Sep 2015), 031044.

[230] VACCARO, J. A., ANSELMI, F., AND WISEMAN, H. M. Entanglement of identical particles and reference phase uncertainty. *International Journal of Quantum Information 1*, 04 (2003), 427–441.

[231] VACCARO, J. A., ANSELMI, F., WISEMAN, H. M., AND JACOBS, K. Tradeoff between extractable mechanical work, accessible entanglement, and ability to act as a reference system, under arbitrary superselection rules. *Physical Review A 77* (Mar 2008), 032114.

[232] VEDRAL, V., AND PLENIO, M. B. Entanglement measures and purification procedures. *Physical Review A 57*, 3 (1998), 1619.

[233] VERSTEEGH, M. A., AND DIEKS, D. The gibbs paradox and the distinguishability of identical particles. *American Journal of Physics 79*, 7 (2011), 741–746.

[234] VINJANAMPATHY, S., AND ANDERS, J. Quantum thermodynamics. *Contemporary Physics 57*, 4 (oct 2016), 545–579.

[235] WANG, X.-B. Theorem for the beam-splitter entangler. *Physical Review A 66*, 2 (aug 2002), 024303.

[236] WANG, Y.-T., TANG, J.-S., WEI, Z.-Y., YU, S., KE, Z.-J., XU, X.-Y., LI, C.-F., AND GUO, G.-C. Directly measuring the degree of quantum coherence using interference fringes. *Physical Review Letters. 118* (Jan 2017), 020403.

[237] WATANABE, G., VENKATESH, B. P., TALKNER, P., HWANG, M.-J., AND DEL CAMPO, A. Quantum statistical enhancement of the collective performance of multiple bosonic engines. *Physical Review Letters 124*, 21 (2020), 210603.

[238] WEI, B.-B., AND PLENIO, M. B. Relations between dissipated work in non-equilibrium process and the family of rényi divergences. *New Journal of Physics. 19* (2017), 023002.

[239] WEI, T.-C. Exchange symmetry and global entanglement and full separability. *Physical Review A 81*, 5 (2010), 054102.

[240] WICK, G. C., WIGHTMAN, A. S., AND WIGNER, E. P. The Intrinsic Parity of Elementary Particles. *Physical Review 88*, 1 (oct 1952), 101–105.

[241] WIGNER, E. *Gruppentheorie.* Vieweg, Braunschweig, 1931.

[242] WINTER, A., AND YANG, D. Operational resource theory of coherence. *Physical Review Letters 116*, 12 (2016), 120404.

[243] WISEMAN, H., AND VACCARO, J. A. Entanglement of indistinguishable particles shared between two parties. *Physical Review Letters 91*, 9 (2003), 097902.

[244] WISEMAN, H. M., BARTLETT, S. D., AND VACCARO, J. A. Ferreting out the fluffy bunnies: Entanglement constrained by generalized superselection rules. In *Laser Spectroscopy.* World Scientific, 2004, pp. 307–314.

[245] WU, K.-D., HOU, Z., ZHONG, H.-S., YUAN, Y., XIANG, G.-Y., LI, C.-F., AND GUO, G.-C. Experimentally obtaining maximal coherence via assisted distillation process. *Optica 4*, 4 (Apr 2017), 454.

[246] XIONG, T., YAN, L., ZHOU, F., REHAN, K., LIANG, D., CHEN, L., YANG, W., MA, Z., FENG, M., AND VEDRAL, V. Experimental verification of a jarzynski-related information-theoretic equality by a single trapped ion. *Physical Review Letters 120*, 1 (2018), 010601.

[247] YADIN, B., BINDER, F. C., THOMPSON, J., NARASIMHACHAR, V., GU, M., AND KIM, M. S. Operational Resource Theory of Continuous-Variable Nonclassicality. *Physical Review X 8*, 4 (dec 2018), 041038.

[248] YADIN, B., MA, J., GIROLAMI, D., GU, M., AND VEDRAL, V. Quantum Processes Which Do Not Use Coherence. *Physical Review X 6*, 4 (nov 2016), 041028.

[249] YADIN, B., MORRIS, B., AND ADESSO, G. Extracting work from mixing indistinguishable systems: A quantum gibbs" paradox". *arXiv preprint arXiv:2006.12482* (2020).

[250] YADIN, B., AND VEDRAL, V. General framework for quantum macroscopicity in terms of coherence. *Physical Review A 93*, 2 (2016), 022122.

[251] YURKE, B., AND STOLER, D. Bell's-inequality experiments using independent-particle sources. *Physical Review A 46*, 5 (sep 1992), 2229–2234.

[252] ZANARDI, P., LIDAR, D. A., AND LLOYD, S. Quantum tensor product structures are observable induced. *Physical Review Letters 92*, 6 (2004), 060402.

[253] ZHAO, Q., LIU, Y., YUAN, X., CHITAMBAR, E., AND MA, X. One-shot coherence dilution. *Physical Review Letters. 120* (Feb 2018), 070403.

[254] ZHU, H., MA, Z., CAO, Z., FEI, S.-M., AND VEDRAL, V. Operational one-to-one mapping between coherence and entanglement measures. *Physical Review A 96* (Sep 2017), 032316.

[255] ZUREK, W. H. Maxwells demon, szilards engine and quantum measurements. In *Frontiers of nonequilibrium statistical physics*. Springer, 1986, pp. 151–161.

[256] ZUREK, W. H. Quantum discord and maxwells demons. *Physical Review A 67*, 1 (2003), 012320.

Appendix A

A.1 Distillable work under thermal operations

Here we present an explicit proof of Eq. (2.3) in the main text, formalizing the connection between the distillable work under TO and the relative entropy of athermality. In what follows, will drop the system subscript B for simplicity.

We start by recalling the main result of [41]: given an initial state ω of a system with Hamiltonian H, and a target state ω' of a system with Hamiltonian H', the asymptotically achievable rates R in a state transformation $\omega^{\otimes n} \to (\omega')^{\otimes [Rn]}$ operated by TO at background inverse temperature β with vanishing error must satisfy

$$S\left(\omega^{\otimes n}\middle\|\gamma^{\otimes n}\right) \geq S\left((\omega')^{\otimes [Rn]}\middle\|(\gamma')^{\otimes [Rn]}\right), \tag{A.1}$$

where $\gamma := Z^{-1}e^{-\beta H}$ and $\gamma' := (Z')^{-1}e^{-\beta H'}$ are the Gibbs states of the input and output systems, respectively. In our case, $\omega^{\otimes n} = \rho^{\otimes n} \otimes |0\rangle\langle0|_P^{\otimes [Rn]}$ and $(\omega')^{\otimes [Rn]} = |1\rangle\langle1|_P^{\otimes [Rn]}$, so that

$$\begin{aligned}
0 &\leq S\left(\rho^{\otimes n} \otimes |0\rangle\langle0|_P^{\otimes [Rn]} \middle\| \gamma^{\otimes n} \otimes \gamma_P^{\otimes [Rn]}\right) - S\left(|1\rangle\langle1|_P^{\otimes [Rn]} \middle\| \gamma_P^{\otimes [Rn]}\right) \\
&= nS(\rho\|\gamma) + [Rn]\, S\left(|0\rangle\langle0|_P \|\gamma_P\right) - [Rn]\, S\left(|1\rangle\langle1|_P \|\gamma_P\right) \\
&= nS(\rho\|\gamma) + [Rn]\log\left(1 + e^{-\beta E}\right) - [Rn]\log\left(\frac{1 + e^{-\beta E}}{e^{-\beta E}}\right) \\
&= nS(\rho\|\gamma) - [Rn]\,\beta E\,,
\end{aligned}$$

where we observed that for the qubit battery one has

$$\gamma_P = \frac{1}{1 + e^{-\beta E}}\begin{pmatrix} 1 & 0 \\ 0 & e^{-\beta E} \end{pmatrix}. \tag{A.2}$$

We deduce that

$$RE \leq \frac{1}{\beta} S(\rho \| \gamma) \,.$$

Taking the supremum as dictated by Eq. (2.2) yields Eq. (2.3), as claimed.

A.2 Bipartite quantum thermal states

As we have seen in the main text, the assisted framework for work distillation allows Alice to perform any given POVM $\{\Pi_{A,i}\}_i$. It is natural to ask, under what conditions the assistance by Alice is a valuable resource that helps to distil more work. In this section we show that this is indeed the case whenever the bipartite state Alice and Bob share is not a product state. In other words, any state which cannot be created for free via the allowed operations constitutes a resource for extracting work on Bob's side.

Proposition 11. *Whenever $\rho_{AB} \neq \rho_A \otimes \rho_B$ is not factorized, we have that*

$$W_a^{B|A}(\rho_{AB}) > \frac{1}{\beta} S(\rho_B \| \gamma_B) \,. \tag{A.3}$$

In other words, the assistance by Alice allows to extract more work than Bob could in the unassisted setting.

Before we present a proof of the above result, it may be useful to recall an elementary lemma.

Lemma 12. *A bipartite quantum state ρ_{AB} is factorized iff for all $\Pi_A \geq 0$ the operator $\mathrm{Tr}_A [\Pi_A \otimes \mathbb{1}_B \, \rho_{AB}]$ is proportional to a fixed state σ_B (independent of Π_A).*

Proof. Since any operator can be written as a complex linear combination of at most four positive operators, we deduce that for all operators N_A we have that $\mathrm{Tr}_A [N_A \otimes \mathbb{1}_B \, \rho_{AB}] = c(N) \, \sigma_B$, where $c(N)$ is a complex scalar. Choosing $N_A = |i\rangle\langle j|_A$ for some basis $|i\rangle_A$ of the Hilbert space on A, and then summing over i, j, one obtains

$$\begin{aligned}
\rho_{AB} &= \sum_{i,j} |i\rangle\langle j|_A \otimes \mathrm{Tr}_A [|i\rangle\langle j|_A \otimes \mathbb{1}_B \, \rho_{AB}] \\
&= \sum_{i,j} |i\rangle\langle j|_A \otimes c_{ij} \sigma_B \\
&= \left(\sum_{i,j} c_{ij} |i\rangle\langle j| \right)_A \otimes \sigma_B \,,
\end{aligned}$$

implying that $\rho_{AB} = \rho_A \otimes \rho_B$ is factorized. $\qquad\square$

We are now ready to prove the main result of this section.

Proof of Proposition 11. Upon measuring her subsystem, Alice will leave Bob in the state $\tilde{\rho}_{B,i} \propto \mathrm{Tr}_A\left[\Pi_{A,i} \otimes \mathbb{1}_B\,\rho_{AB}\right]$ of (2.4) with probability $p_i = \mathrm{Tr}\left[\rho_A \Pi_{A,i}\right]$. By Lemma 12, we know that if $\rho_{AB} \neq \rho_A \otimes \rho_B$ there are two positive operators $\Pi_{A,1}, \Pi_{A,2} \geq 0$ such that $\tilde{\rho}_{B,1} \neq \tilde{\rho}_{B,2}$. Up to rescaling them, we can make sure that $\Pi_{A,1} + \Pi_{A,2} \leq \mathbb{1}_A$, so that there exists a valid POVM that includes both $\Pi_{A,1}$ and $\Pi_{A,2}$. Invoking the strict concavity of the entropy, it is then elementary to establish that

$$
\begin{aligned}
\beta\,W_a^{B|A}(\rho_{AB}) &\geq \sum_i p_i S(\tilde{\rho}_{B,i}\|\gamma_B) \\
&= -\sum_i p_i S(\tilde{\rho}_{B,i}) - \sum_i p_i \,\mathrm{Tr}\left[\tilde{\rho}_{B,i} \log \gamma_B\right] \\
&> -S\left(\sum_i p_i \tilde{\rho}_{B,i}\right) - \mathrm{Tr}\left[\left(\sum_i p_i \tilde{\rho}_{B,i}\right) \log \gamma_B\right] \\
&= -S(\rho_B) - \mathrm{Tr}\left[\rho_B \log \gamma_B\right] \\
&= S(\rho_B\|\gamma_B)\,,
\end{aligned}
$$

which concludes the proof. $\square$

A.3   Work of assistance as a function of the Henderson–Vedral measure

Suppose Alice performs a general quantum operation $\{\Pi_{A,i}\}_i$, causing Bob to receive a state $\tilde{\rho}_{B,i}$, with the associated probabilities $\mathrm{Tr}\,\rho_A \Pi_{A,i} = p_i$. She can use her classical communication channel to choose a measurement so that Bob's extractable work is maximized,

$$
W_a^{B|A}(\rho_{AB}) = \max_{\{\Pi_{A,i}\}} \frac{1}{\beta} \sum_i p_i S\left(\tilde{\rho}_{B,i}\|\gamma_B\right), \tag{A.4}
$$

$$
= \max_{\{\Pi_{A,i}\}} \left(\langle E\rangle_{\tilde{\rho}_{B,i}} + \frac{1}{\beta}\left[\ln Z_B - \sum_i p_i S(\tilde{\rho}_{B,i})\right]\right). \tag{A.5}
$$

The work of assistance can also be written in the following way,

$$
\begin{aligned}
W_a^{B|A}(\rho_{AB}) &= \max_{\{\Pi_{A,i}\}} \frac{1}{\beta} \sum_i p_i S(\tilde{\rho}_{B,i} || \gamma_B) \\
&= \max_{\{\Pi_{A,i}\}} \frac{1}{\beta} \sum_i p_i \operatorname{Tr}\left[\tilde{\rho}_{B,i} \left(\log \tilde{\rho}_{B,i} - \log \gamma_B \right) \right] \\
&= \max_{\{\Pi_{A,i}\}} \frac{1}{\beta} \sum_i p_i \left[-S(\tilde{\rho}_{B,i}) - \operatorname{Tr} \tilde{\rho}_{B,i} \log \gamma_B \right] \\
&= \max_{\{\Pi_{A,i}\}} \frac{1}{\beta} \left[- \sum_i p_i S(\tilde{\rho}_{B,i}) - \operatorname{Tr} \rho_B \log \gamma_B \right],
\end{aligned}
\tag{A.6}
$$

as $\sum p_i \tilde{\rho}_{B,i} = \rho_B$. Therefore the work of assistance refers to Alice attempting to minimize the local entropy on Bob's side via her measurement,

$$
W_a^{B|A}(\rho_{AB}) = -\frac{1}{\beta} \left[\operatorname{Tr} \rho_B \log \gamma_B + \min_{\{\Pi_{A,i}\}_i} \sum_i p_i S(\tilde{\rho}_{B,i}) \right].
\tag{A.7}
$$

We can lower bound this value using the convexity of the quantum relative entropy

$$
W_a^{B|A}(\rho_{AB}) \geq \frac{1}{\beta} S(\rho_B || \gamma_B),
\tag{A.8}
$$

which is also additive. In order to further investigate the work of assistance we can employ a measure of classical correlations from [117] defined as

$$
J^\rightarrow(\rho_{AB}) = \max_{\{\Pi_{A,i}\}_i} \left[S(\rho_B) - \sum_i p_i S(\tilde{\rho}_{B,i}) \right],
\tag{A.9}
$$

where, like the work of assistance (A.6), the maximization is taken over all the measurements $\{\Pi_{A,i}\}$ applied on Alice's subsystem. Substituting this into (2.5) allows us to write the work of assistance as

$$
\begin{aligned}
W_a^{B|A}(\rho_{AB}) &= -\frac{1}{\beta} \operatorname{Tr}\left[\rho_B \log \gamma_B \right] - \frac{1}{\beta} S(\rho_B) + \frac{1}{\beta} J^\rightarrow(\rho_{AB}), \\
&= \frac{1}{\beta} \left(S(\rho_B || \gamma_B) + J^\rightarrow(\rho_{AB}) \right).
\end{aligned}
\tag{A.10}
$$

A.4 The regularized work of assistance is the maximum distillable work with one-way communication

Here we argue that the regularized work of assistance $W_{a,\infty}^{B|A}$ of Eq. (2.7) coincides with the optimal work distillation rate when only Alice $\rightarrow$ Bob classical communication is allowed. In this case, it makes no difference whether Bob has access to all Gibbs-preserving operations or only to thermal ones. Therefore, our claim also implies that the work of collaboration $W_c^{B|A}$ of Eq. (2.15) coincides with $W_{a,\infty}^{B|A}$ when the operations on Bob's side are required to be thermal, i.e. $\mathcal{O} = \mathrm{TO}$.

Start by considering a rate R that is achievable for a fixed state ρ of AB in the sense of Eq. (2.15), where we assume from now on that $\mathcal{O} = \mathrm{TO}$. This means that there is a sequence $\{\Lambda_n\}_n$ of operations on $A : BP$ obtained by concatenating arbitrary quantum instruments on Alice, classical communication Alice $\rightarrow$ Bob, and TO on Bob's side (which includes the battery P), such that

$$\Lambda_n\left(\rho_{AB}^{\otimes n} \otimes |0\rangle\langle 0|_P^{\otimes Rn}\right) = |1\rangle\langle 1|_P^{\otimes [Rn]} + \delta_n\,,$$

where the above equation defines a sequence of "remainder terms" $\{\delta_n\}_n$ that: (i) are traceless, i.e. such that $\mathrm{Tr}\,\delta_n = 0$; (ii) satisfy $\epsilon_n := \|\delta_n\|_1 \rightarrow 0$ as $n \rightarrow \infty$. By definition of work of assistance, we can write

$$
\begin{aligned}
\frac{1}{n}W_a^{B|A}\left(\rho_{AB}^{\otimes n}\right) &\geq \frac{1}{n\beta}S\left(|1\rangle\langle 1|_P^{\otimes [Rn]} + \delta_n \,\Big\|\, \gamma_P^{\otimes n}\right) \\
&= -\frac{1}{n\beta}S\left(|1\rangle\langle 1|_P^{\otimes [Rn]} + \delta_n\right) - \frac{1}{n\beta}\mathrm{Tr}\left[\left(|1\rangle\langle 1|_P^{\otimes [Rn]} + \delta_n\right)\log\left(\gamma_P^{\otimes [Rn]}\right)\right] \\
&= -\frac{1}{n\beta}S\left(|1\rangle\langle 1|_P^{\otimes [Rn]} + \delta_n\right) - \frac{[Rn]}{n\beta}\log\left(\frac{e^{-\beta E}}{1 + e^{-\beta E}}\right) - \frac{1}{n\beta}\mathrm{Tr}\left[\delta_n \log\left(\gamma_P^{\otimes [Rn]}\right)\right] \\
&\geq -\frac{1}{n\beta}S\left(|1\rangle\langle 1|_P^{\otimes [Rn]} + \delta_n\right) + \frac{[Rn]}{n}E - \frac{1}{n\beta}\mathrm{Tr}\left[\delta_n \log\left(\gamma_P^{\otimes [Rn]}\right)\right]\,,
\end{aligned}
$$
$$(A.11)$$

where we remembered that the Gibbs state of the qubit battery is given by Eq. (A.2). Rearranging and taking the limit $n \rightarrow \infty$, we obtain

$$RE = \lim_{n\to\infty}\frac{[Rn]}{n}E \leq W_{a,\infty}^{B|A}(\rho_{AB}) + \lim_{n\to\infty}\frac{1}{n\beta}S\left(|1\rangle\langle 1|_P^{\otimes [Rn]} + \delta_n\right) + \lim_{n\to\infty}\frac{1}{n\beta}\mathrm{Tr}\left[\delta_n \log\left(\gamma_P^{\otimes [Rn]}\right)\right]\,,$$
$$(A.12)$$

provided that those limits exist. We now claim that the second and third term on the

r.h.s. of Eq. (A.12) vanish. Indeed, by Fannes's inequality [83, 14, 242] one can write

$$
\begin{aligned}
\frac{1}{n} S \left(|1\rangle\langle 1|_P^{\otimes [Rn]} + \delta_n \right) &\leq \frac{1}{n} \left(S \left(|1\rangle\langle 1|_P^{\otimes [Rn]} \right) + \frac{1}{2}\epsilon_n \log \left(2^{[Rn]} \right) + h_2 \left(\epsilon_n/2 \right) \right) \\
&= \frac{[Rn]}{2n} \epsilon_n + \frac{1}{n} h_2 \left(\epsilon_n/2 \right) \\
&\xrightarrow[n\to\infty]{} 0 \,,
\end{aligned}
\tag{A.13}
$$

where in the last step we used the fact that $\lim_{n\to\infty} \epsilon_n = 0$. As for the third term on the r.h.s. of Eq. (A.12), we observe that

$$
\log \left(\gamma_P^{\otimes [Rn]} \right) = (\log \gamma_P) \otimes \mathbb{1} \otimes \ldots \otimes \mathbb{1} + \mathbb{1} \otimes (\log \gamma_P) \otimes \ldots \otimes \mathbb{1} + \ldots + \mathbb{1} \otimes \ldots \otimes \mathbb{1} \otimes (\log \gamma_P) \,,
$$

from which one deduces immediately that

$$
\left\| \log \left(\gamma_P^{\otimes [Rn]} \right) \right\|_\infty = [Rn] \left\| \log \gamma_P \right\|_\infty \,.
$$

Applying Hölder's inequality, we then obtain

$$
\begin{aligned}
\frac{1}{n} \operatorname{Tr} \left[\delta_n \log \left(\gamma_P^{\otimes [Rn]} \right) \right] &\leq \frac{1}{n} \|\delta_n\|_1 \left\| \log \left(\gamma_P^{\otimes [Rn]} \right) \right\|_\infty \\
&= \frac{[Rn]}{n} \epsilon_n \| \log \gamma_P \|_\infty \\
&\xrightarrow[n\to\infty]{} 0 \,,
\end{aligned}
\tag{A.14}
$$

where the last step is made possible by the fact that $\| \log \gamma_P \|_\infty$ is a constant independent of n.

We have thus established that the r.h.s. of Eq. (A.12) coincides with the regularized work of assistance $W_{a,\infty}$, which then upper bounds any achievable rate in the expression for the work of collaboration corresponding to the setting where Bob can only access TO. This concludes the proof.

A.5   Work of assistance for isotropic states

We now continue with an explicit example of how the work of assistance can be computed by considering the following non-trivial family of bipartite states. The isotropic states [125] on a $d \times d$ system appear as

$$
\rho_{AB}(\lambda) := \lambda \Phi_d + \frac{(1-\lambda)}{d^2} \mathbb{1}_{AB},
\tag{A.15}
$$

where $\Phi_d = |\Phi_d\rangle\langle\Phi_d|$ is the maximally entangled state $|\Phi_d\rangle := \frac{1}{\sqrt{d}}\sum_{i=1}^{d}|ii\rangle$. The isotropic state is a legitimate state ($\rho_{AB}(\lambda) \geq 0$) for $-\frac{1}{d^2-1} \leq \lambda \leq 1$ and entangled for $\lambda > \frac{1}{d+1}$. The choice of an isotropic state allows us to quantify the role of entanglement in the work of assistance through the constant λ.

In this section we will attempt to determine the ensemble $\{p_i, \tilde{\rho}_{B,i}\}$ of d-dimensional states obtained when Alice measures her subsystem of an isotropic state $\rho_{AB}(\lambda)$ where λ is a priori fixed. We will then attempt to determine the form of this ensemble that maximizes the distillable work on Bob's subsystem.

Consider a POVM $\{\Pi_{A,i}\}$ such that $p_i = \mathrm{Tr}_{AB}[(\Pi_{A,i} \otimes \mathbb{1})\rho_{AB}(\lambda)]$ and $p_i\tilde{\rho}_i = \mathrm{Tr}_A[(\Pi_{A,i} \otimes \mathbb{1})\rho_{AB}(\lambda)]$. The average of this ensemble can be calculated as

$$
\begin{aligned}
\sum_i p_i\tilde{\rho}_i &= \mathrm{Tr}_A\left[\left(\sum_i \Pi_{A,i} \otimes \mathbb{1}_B\right)\rho_{AB}(\lambda)\right], \\
&= \mathrm{Tr}_A\left[\rho_{AB}(\lambda)\right], \\
&= \rho_B(\lambda), \\
&= \frac{\mathbb{1}_B}{d} \quad \forall \ \lambda.
\end{aligned}
\tag{A.16}
$$

The states in the ensemble appear as,

$$
\begin{aligned}
\tilde{\rho}_i &= \frac{1}{p_i}\mathrm{Tr}_A\left[(\Pi_{A,i} \otimes \mathbb{1}_B)\left(\lambda\Phi_d + \frac{(1-\lambda)}{d^2}\mathbb{1}_{AB}\right)\right], \\
&= \frac{1}{p_i}\left\{\lambda\,\mathrm{Tr}_A\left[\left(\sqrt{\Pi_{A,i}} \otimes \mathbb{1}_B\right)|\Phi_d\rangle\langle\Phi_d|\left(\sqrt{\Pi_{A,i}} \otimes \mathbb{1}_B\right)\right] + \frac{1-\lambda}{d^2}\mathrm{Tr}_A\left[(\Pi_{A,i} \otimes \mathbb{1}_B)\mathbb{1}_{AB}\right]\right\},
\end{aligned}
\tag{A.17}
$$

Using the identity $M \otimes \mathbb{1}|\Psi\rangle = \mathbb{1} \otimes M^\dagger|\Psi\rangle$, we have

$$
\begin{aligned}
\tilde{\rho}_i &= \frac{1}{p_i}\left\{\lambda\,\mathrm{Tr}_A\left[\left(\mathbb{1}_A \otimes \sqrt{(\Pi_{B,i})^T}\right)|\Phi_d\rangle\langle\Phi_d|\left(\mathbb{1}_A \otimes \sqrt{(\Pi_{B_i})^T}\right)\right] + \frac{1-\lambda}{d^2}\mathrm{Tr}\left[\Pi_i\right]\mathbb{1}_B\right\}, \\
&= \frac{1}{p_i}\left\{\lambda\sqrt{(\Pi_{B,i})^T}\frac{\mathbb{1}_B}{d}\sqrt{(\Pi_{B,i})^T} + \frac{1-\lambda}{d^2}\mathrm{Tr}[\Pi_i]\mathbb{1}_B\right\}, \\
&= \frac{1}{p_i}\left\{\frac{\lambda}{d}(\Pi_{B,i})^T + \frac{1-\lambda}{d^2}\mathrm{Tr}[\Pi_i]\mathbb{1}_B\right\}.
\end{aligned}
\tag{A.18}
$$

By imposing $\operatorname{Tr} \tilde{\rho}_i = 1$, this yields,

$$
\begin{aligned}
1 &= \operatorname{Tr} \tilde{\rho}_i, \\
&= \frac{1}{p_i} \left\{ \frac{\lambda}{d} \operatorname{Tr} \Pi_i + \frac{1 - \lambda}{d} \operatorname{Tr} \Pi_i \right\}, \\
&= \frac{1}{p_i} \frac{\operatorname{Tr} \Pi_i}{d}.
\end{aligned}
\tag{A.19}
$$

Therefore,

$$
p_i = \frac{\operatorname{Tr} \Pi_i}{d}
\tag{A.20}
$$

and hence,

$$
\tilde{\rho}_i = \lambda \frac{(\Pi_{B,i})^T}{\operatorname{Tr} \Pi_i} + \frac{1 - \lambda}{d} \mathbb{1}_B.
\tag{A.21}
$$

Depending on whether our fixed variable λ is positive or negative, as $\Pi_{B,i} \geq 0$ we can write the following,

$$
\tilde{\rho}_i \begin{cases} \geq \frac{1-\lambda}{d} \mathbb{1}_B \text{ for } \lambda \geq 0 \\[2em] < \frac{1-\lambda}{d} \mathbb{1}_B \text{ for } \lambda < 0 \end{cases}.
\tag{A.22}
$$

Which means we can write the ensembles states as

$$
\tilde{\rho}_i = \frac{1 - \lambda}{d} \mathbb{1} \pm \delta_i, \quad \delta_i \geq 0 \ \forall i.
\tag{A.23}
$$

From this ensemble we will attempt to find a POVM on A that generates it. We start from the fact that the average of this ensemble is the maximally mixed state,

$$
\begin{aligned}
\frac{\mathbb{1}}{d} &= \sum_i p_i \tilde{\rho}_i, \\
&= \sum_i p_i \left(\frac{1 - \lambda}{d} \mathbb{1} \pm \delta_i \right), \\
&= \frac{1 - \lambda}{d} \mathbb{1} \pm \sum_i p_i \delta_i.
\end{aligned}
\tag{A.24}
$$

From this we can construct our set of POVM operators

$$\pm \sum_i p_i \delta_i = \lambda \frac{\mathbb{1}}{d},$$

$$\implies \quad \pm \sum_i \frac{d}{\lambda} p_i \delta_i = \mathbb{1}, \tag{A.25}$$

implying that the above operators form a valid POVM if they have the form $\Pi_i := \pm \frac{d}{\lambda} p_i \delta_i$.

This POVM can be checked with our current ensemble,

$$\left\{ p_i = \frac{\operatorname{Tr} \Pi_i}{d}, \quad \tilde{\rho}_i = \lambda \frac{(\Pi_i)^T}{\operatorname{Tr} \Pi_i} + \frac{1-\lambda}{d} \mathbb{1} \right\}_i, \tag{A.26}$$

in our case,

$$\begin{aligned}
p_i &= \frac{\operatorname{Tr} \Pi_i}{d} = \pm \frac{1}{d} \operatorname{Tr} \frac{d}{\lambda} p_i \delta_i = \pm p_i \frac{\operatorname{Tr} \delta_i}{\lambda}, \\
&= p_i \frac{\operatorname{Tr} \left[\tilde{\rho}_i - \frac{1-\lambda}{d} \mathbb{1} \right]}{\lambda} = p_i \frac{1 - (1-\lambda)}{\lambda}, \\
&= p_i, \tag{A.27}
\end{aligned}$$

and

$$\begin{aligned}
\tilde{\rho}_i &= \lambda \frac{(\Pi_i)^T}{\operatorname{Tr} \Pi_i} + \frac{1-\lambda}{d} \mathbb{1} = \lambda \frac{\pm \frac{d}{\lambda} p_i \delta_i}{\pm \frac{d}{\lambda} p_i \operatorname{Tr} \delta_i} + \frac{1-\lambda}{d} \mathbb{1}, \\
&= \lambda \frac{\pm \delta_i}{\pm \operatorname{Tr} \delta_i} + \frac{1-\lambda}{d} \mathbb{1} = \lambda \frac{\tilde{\rho}_i - \frac{1-\lambda}{d} \mathbb{1}}{\lambda} + \frac{1-\lambda}{d} \mathbb{1}, \\
&= \tilde{\rho}_i - \frac{1-\lambda}{d} \mathbb{1} + \frac{1-\lambda}{d} \mathbb{1}, \\
&= \tilde{\rho}_i. \tag{A.28}
\end{aligned}$$

confirming our choice of POVM.

In order to use this conceived ensemble in the work of assistance (2.5) it must minimize the entropy. In order for a mixed state to minimize its entropy it must concentrate its spectrum on a single eigenvalue, making it as pure as possible. The conditions on these states (A.21) and probabilities (A.20) are that their average is the maximally mixed state and that equation (A.22) is satisfied for positive or negative fixed values of λ.

An ensemble for which all the above holds is the following,

$$\left\{ p_i = \frac{1}{d}, \ \ \tilde{\rho}_i = \frac{1-\lambda}{d}\mathbb{1} + \lambda|i\rangle\langle i| \ \forall i \right\}_i. \tag{A.29}$$

which, using $\lambda + \frac{1-\lambda}{d} = 1 - (d-1)\frac{1-\lambda}{d}$, allows us to write down the work of assistance for this family of states,

$$W_a^{B|A}(\rho_{AB}(\lambda)) = -\frac{1}{\beta}\left[\mathrm{Tr}\,\frac{1}{d}\log\gamma_B + H\left(\overbrace{\frac{1-\lambda}{d}, ..., \frac{1-\lambda}{d}}^{d-1}, 1 - (d-1)\frac{1-\lambda}{d}\right)\right], \tag{A.30}$$

where $H(p_1, ..., p_d) = -\sum_i p_i \log p_i$ is the Shannon entropy.

It is known [142] [115, p. 448] that the Henderson-Vedral measure $J^{\rightarrow}$ and hence the work of assistance (2.6) is additive over the family of isotropic states.

A.6 Monotonicity and rewriting of $W_r^{B|A}(\rho_{AB})$

We start with the following definition,

$$W_r^{B|A}(\rho_{AB}) = \frac{1}{\beta}\min_{\Gamma_{AB}\in QT} S(\rho_{AB}||\Gamma_{AB}), \tag{A.31}$$

where the minimization is taken over the set QT.

As the thermal state γ_B appearing within the QT state $\Gamma_{AB} = \sigma_A \otimes \gamma_B$ is fixed, the minimization should be taken over Alice's state σ_A,

$$
\begin{aligned}
W_r^{B|A}(\rho_{AB}) &= \frac{1}{\beta}\min_{\sigma_A} S(\rho_{AB}||\sigma_A \otimes \gamma_B) \\
&= \frac{1}{\beta}\min_{\sigma_A}\left\{-\mathrm{Tr}\,\rho_{AB}(\log\sigma_A + \log\gamma_B) - S(\rho_{AB})\right\} \\
&= \frac{1}{\beta}\left(-\left(\max_{\sigma_A}\mathrm{Tr}\,\rho_A\log\sigma_A\right) - \mathrm{Tr}\,\rho_B\log\gamma_B - S(\rho_{AB})\right).
\end{aligned} \tag{A.32}
$$

Now using the following, $S(\rho_A) = -\max_{\sigma_A}\mathrm{Tr}\,\rho_A\log\sigma_A$, as $\min_{\sigma_A}S(\rho_A||\sigma_A) = 0$, in (A.32), we get

$$
\begin{aligned}
W_r^{B|A}(\rho_{AB}) &= \frac{1}{\beta}\left(-S(\rho_{AB}) + S(\rho_A) - \mathrm{Tr}\,\rho_B\log\gamma_B\right) \\
&= \frac{1}{\beta}S(\rho_{AB}||\rho_A \otimes \gamma_B).
\end{aligned} \tag{A.33}
$$

If we take the bipartite quantum state $\Lambda(\rho_{AB}) = \tilde{\rho}_{A'B}$ where Λ are the set of allowed operations, we can write,

$$
\begin{aligned}
W_r^{B|A}(\tilde{\rho}_{A'B}) =& \frac{1}{\beta} \inf_{\tilde{\sigma}_{A'}} S\left(\tilde{\rho}_{A'B} || \tilde{\sigma}_{A'} \otimes \gamma_B\right) \\
\leq& \frac{1}{\beta} S\left(\tilde{\rho}_{A'B} || \rho_A \otimes \gamma_B\right) \\
\leq& \frac{1}{\beta} S\left(\rho_{AB} || \rho_A \otimes \gamma_B\right) \\
=& W_r^{B|A}(\rho_{AB}),
\end{aligned}
\tag{A.34}
$$

where $\Lambda\left(\rho_A \otimes \gamma_B\right) = \tilde{\sigma}_{A'} \otimes \gamma_B$, demonstrating the monotonicity of $W_r^{B|A}(\rho_{AB})$ under the set of allowed operations. Also due to the additivity property of relative entropy, $W_r^{B|A}(\rho_{AB})$ is also additive.

We also note that $W_r^{B|A}(\rho_{AB})$ can be expressed in relation to the quantum mutual information,

$$
\begin{aligned}
W_r^{B|A}(\rho_{AB}) =& \frac{1}{\beta} \left(-\operatorname{Tr} \rho_{AB} \log(\rho_A \otimes \gamma_B) - S(\rho_{AB})\right) \\
=& \frac{1}{\beta} \left(-\operatorname{Tr} \rho^A \log(\rho_A) - \operatorname{Tr} \rho_B \log(\gamma_B) - S(\rho_{AB})\right) \\
=& \frac{1}{\beta} \left(S(\rho^A || \rho_A) + S(\rho_B || \gamma_B) + I(\rho_{AB})\right), \\
=& \frac{1}{\beta} \left(S(\rho_B || \gamma_B) + I(\rho_{AB})\right).
\end{aligned}
\tag{A.35}
$$

which is a measure of the shared total correlations between the two parties.

A.7   Upper bound on the work of collaboration

In this section we will prove that the work of collaboration is upper bounded by $W_r^{B|A}(\rho_{AB})$,

$$
W_r^{B|A}(\rho_{AB}) \geq W_c^{B|A}(\rho_{AB}).
\tag{A.36}
$$

This proof will follow analogously to [232, 58]. Let R be an achievable rate for Eq. (2.15), so that there exists a sequence of protocols $\Lambda_n \in \mathcal{O}_c^{B|A}$ with the property that $\Lambda_n\left(\rho_{AB}^{\otimes n} \otimes |0\rangle\langle 0|_P^{\otimes [Rn]}\right) = |1\rangle\langle 1|_P^{\otimes [Rn]} + \delta_n$, where $\lim_{n\to\infty} \|\delta_n\|_1 = 0$. Using the monotonicity of $W_r^{B|A}$ together with its additivity, and repeating the steps in Eq. (A.11), we obtain

that

$$
\begin{aligned}
W_r^{B|A}(\rho_{AB}) &= \frac{1}{n} W_r^{B|A} \left(\rho_{AB}^{\otimes n} \otimes |0\rangle\langle 0|_P^{\otimes [Rn]} \right) \\
&\geq \frac{1}{n} W_r^{B|A} \left(|1\rangle\langle 1|_P^{\otimes [Rn]} + \delta_n \right) \\
&= \frac{1}{\beta n} S \left(|1\rangle\langle 1|_P^{\otimes [Rn]} + \delta_n \middle\| \gamma_P^{\otimes [Rn]} \right) \\
&\geq -\frac{1}{\beta n} S \left(|1\rangle\langle 1|_P^{\otimes [Rn]} + \delta_n \right) + \frac{[Rn]}{n} E - \frac{1}{n\beta} \operatorname{Tr} \left[\delta_n \log \left(\gamma_P^{\otimes [Rn]} \right) \right].
\end{aligned}
$$

Note that the equality in the third line holds because the battery P is on Bob's side. Employing the already established Eq. (A.13) and (A.14), we arrive at the bound $RE \leq W_r^{B|A}(\rho_{AB})$, concluding the proof.

Appendix B

B.1 Proof of necessary and sufficient conditions for B(S)IO state transformations

Here we complete the proof of the Theorem in Section 3.4.

B.1.1 Sufficient conditions

Starting with the conditional probability distribution $P(i, w|j)$ satisfying Conditions 1–3 as discussed in Section 3.4, let us define a sequence of initial and final states of the d-dimensional system A and the battery B, indexed by an even N, as follows,

$$|\Psi^{(N)}\rangle_{AB} = \sum_{i=1}^{d} \sum_{x=0}^{n} \sqrt{\sum_{w} \frac{p_{i,w}}{N+1}} \beth_{x+f_w}(\mathcal{S}) \, |i\rangle_A \otimes |c_x\rangle_B, \tag{B.1}$$

$$|\Phi^{(N)}\rangle_{AB} = \sum_{i=1}^{d} \sum_{x'=0}^{n} \sqrt{\sum_{w} \frac{q_j}{N+1}} \beth_{x'}(\mathcal{S}) \, |j\rangle_A \otimes |c_{x'}\rangle_B. \tag{B.2}$$

Here $p_{i,w} := \sum_{j=1}^{d} P(i, w|j) q_j$ and the integer factor

$$f_w := \left\lfloor \frac{w}{\delta w} \right\rfloor \tag{B.3}$$

measures the fluctuating coherence w in units of δw, while the symbol $\beth_k(\mathcal{S})$ with $\mathcal{S} := \left\{ \frac{n-N}{2}, ..., \frac{n+N}{2} \right\}$ denotes an indicator function, i.e., it equals 1 if the index $k \in \mathcal{S}$ and 0 otherwise. Furthermore, we set $N := n - 2f_{\max}$, with $f_{\max} := \max_{w}\{|f_w|\}$.

We then need to construct a sequence of bistochastic matrices $G^{(N)}$ mapping the

(nonzero) diagonal entries of $\Phi_{AB}^{(N)}$ to those of $\Psi_{AB}^{(N)}$ [37]. The following adapts the proof reported in [9] for entanglement theory, to which the reader is referred for further details.

Let us begin by defining a new probability distribution

$$P(i, x | j, x') := \sum_w P(i, w | j) \delta_{x'-x, f_w}, \tag{B.4}$$

with $x, x' \in (-\infty, \infty)$. The three conditions (3.25)–(3.27) can then be rewritten as

$$\sum_{i=1}^{d} \sum_{x=-\infty}^{\infty} P(i, x | j, x') = 1, \tag{B.5}$$

$$\sum_{i=1}^{d} \sum_{x=-\infty}^{\infty} P(i, x | j, x') \left(\frac{u}{u-1}\right)^{x'-x} \leq 1, \tag{B.6}$$

$$\sum_{j=1}^{d} \sum_{x'=-\infty}^{\infty} P(i, x | j, x') \frac{q_j}{N+1} \beth_{x'}(\mathcal{S}) = \sum_w \frac{p_{i,w}}{N+1} \beth_{x+f_w}(\mathcal{S}), \tag{B.7}$$

where for the second equation we have used the fact that f_w is the largest integer such that $f_w \ln\left(\frac{u}{u-1}\right) \leq w$, which follows from the definitions (3.16) and (B.3).

With this newly defined conditional probability $P(i, x | j, x')$, a sequence of sub-bistochastic matrices can now be constructed, whose rows and columns are labelled by the reference product basis of system and battery, say $|i, z\rangle_{AB}$ and $|j, z'\rangle_{AB}$, respectively. Here and in the following we employ the shorthand notation $z \in \chi_x$ to mean $|z\rangle \in \text{supp}(\chi_x)$. Then, for all $z \in \chi_x$, $z' \in \chi_{x'}$, where now the values of x, x' are truncated as $x, x' \in \{0, ..., n\}$, let

$$R^{(N)}(i, z | j, z') := P(i, x | j, x') u^{-x} (u-1)^{x-n}. \tag{B.8}$$

Using Eqs. (B.5)–(B.7), we see that the matrices $R^{(N)}(i, z | j, z')$ are not bistochastic, but they already possess the desired property of mapping the entries of $\Delta(\Phi^{(N)})$ to those of $\Delta(\Psi^{(N)})$,

$$\sum_{j=1}^{d} \sum_{x'=\frac{n-N}{2}-f_{\max}}^{\frac{n+N}{2}+f_{\max}} \sum_{z \in \chi_{x'}} R(i, z | j, z') \frac{q_j}{N+1} u^{-x'} (u-1)^{x'-n} \beth_{x'}(\mathcal{S})$$

$$= \sum_{j=1}^{d} \sum_{x'=-\infty}^{\infty} P(i, x | j, x') \frac{q_j}{N+1} u^{-x} (u-x)^{x-n} \beth_{x'}(\mathcal{S})$$

$$= \sum_w \frac{p_{i,w}}{N+1} u^{-x} (u-x)^{x-n} \beth_{x+f_w}(\mathcal{S}). \tag{B.9}$$

Finally, we can define a sequence of bistochastic matrices $G^{(N)}(i, z|j, z')$ based on $R^{(N)}(i, z|j, z')$, such that the action of the latter on the support of $\Delta(\Phi^{(N)})$ is left intact. This can be done as follows:

$$G^{(N)}(i, z|j, z') := \begin{cases} R^{(N)}(i, z|j, z'), & \forall \, z' \in \chi_{x'} \text{ with } x' \in \{\frac{n-N}{2}, ..., \frac{n+N}{2}\}; \\ \frac{1}{d\mu}\left(1 - \sum_{j=1}^{d}\sum_{x'=\frac{n-N}{2}}^{\frac{n+N}{2}}\sum_{z'\in\chi_{x'}} R^{(N)}(i, z|j, z')\right), & \text{otherwise}, \end{cases}$$

(B.10)

where $\mu := \sum_{x'=0}^{\frac{n-N}{2}} u^{x'}(u-1)^{n-x'} + \sum_{x'=\frac{n+N}{2}}^{n} u^{x'}(u-1)^{n-x'}$, so that $d\mu$ is the number of columns of $G^{(N)}$ not belonging to the support of $\Delta(\Phi_N)$. It follows by construction that the matrices $G^{(N)}(i, z|j, z')$ are bistochastic,

$$\sum_{i=1}^{d}\sum_{x=0}^{n}\sum_{z\in\chi_x} G^{(N)}(i, z|j, z') = \sum_{j=1}^{d}\sum_{x'=0}^{n}\sum_{z'\in\chi'_x} G^{(N)}(i, z|j, z') = 1\,.$$

(B.11)

Moreover, by virtue of Eq. (B.9), these matrices map the diagonal coefficients of $\left|\Phi^{(N)}\right\rangle_{AB}$ to those of $\left|\Psi^{(N)}\right\rangle_{AB}$. This means that, according to [37], there exists a sequence of (S)IO protocols $\Lambda_{AB}^{(N)}$ that implement the state transformations $\Phi_{AB}^{(N)} = \Lambda_{AB}^{(N)}(\Psi_{AB}^{(N)})$.

We are only left to verify that $\lim_{N\to\infty}\left|\Psi^{(N)}\right\rangle_{AB} = |\psi\rangle_A \otimes |\lambda\rangle_B$ and $\lim_{N\to\infty}\left|\Phi^{(N)}\right\rangle_{AB} = |\phi\rangle_A \otimes |\lambda\rangle_B$. For the final states $\left|\Phi^{(N)}\right\rangle_{AB}$, which are uncorrelated across the system versus battery split for any finite N, this is manifestly true. For the initial states, $\left|\Psi^{(N)}\right\rangle_{AB}$, which are instead entangled across such a split, we need to analyse the large N limit explicitly. It is straightforward to show [9] that the overlaps between the states $\left|\Psi^{(N)}\right\rangle_{AB}$ of Eq. (B.1) and the product states

$$|\widetilde{\Psi}^{(N)}\rangle_{AB} = \sum_{i=1}^{d}\sum_{x=0}^{n}\sqrt{\frac{p_i}{n+1}}|i\rangle_A \otimes |c_x\rangle_B,$$

(B.12)

which reproduce the correct initial state in the limit $N \to \infty$, satisfy

$$\langle\widetilde{\Psi}^{(N)}|\Psi^{(N)}\rangle = \sum_{i=1}^{d}\sum_{x=0}^{n}\sqrt{\frac{p_i}{n+1}}\sum_{w}\frac{p_{i,w}}{N+1}\rrbracket_{x+f_w}(\mathcal{S})$$

$$\geq \sum_{i=1}^{d}\sum_{x=\frac{n-N}{2}+f_{\max}}^{\frac{n+N}{2}-f_{\max}}\sqrt{\frac{p_i^2}{(n+1)(N+1)}}$$

$$\geq \frac{1}{1+\frac{2f_{\max}}{N+1}} \xrightarrow{N\to\infty} 1\,.$$

(B.13)

This concludes the proof of the sufficient conditions for the B(S)IO state transformation (3.1).

B.1.2   Necessary conditions

Here we show that the conditional probability distribution $P(i,w|j)$ defined in Eq. (3.29) satisfies the three conditions detailed in Eqs. (3.25), (3.26), and (3.27).

Before proceeding, let us note that all the (diagonal) battery states considered in the protocol of Section 3.4 live entirely in the subspace $\mathcal{B}$ introduced in Section 3.3. Accordingly, the supports of $\Delta(\Psi_{AB})$ and $\Delta(\Phi_{AB})$ are contained in the relevant subspace $\mathcal{V} := \mathcal{D}(\mathcal{H}_A) \otimes \mathcal{B}$ of the state space of system and battery, and the whole protocol described in Section 3.4 to construct $P(i,w|j)$ maps this subspace into itself. It follows that, as $\Delta(\Psi_{AB}) \prec \Delta(\Phi_{AB})$ by hypothesis, then one can find permutations Ξ_m and probabilities r_m such that the unital map E_{AB} of the form (3.5) satisfies (3.28) and furthermore leaves invariant the projector $\mathbb{1}_A \otimes \mathbb{B}_B$ onto $\mathcal{V}$.

To prove Condition 1 (normalisation), we sum over the initial reference basis i and the extracted coherence w,

$$\sum_{i,w} P(i,w|j) = \sum_{x'} \mathrm{Tr}\left[(\mathbb{1}_A \otimes \mathbb{B}_B)\mathcal{E}_{AB}(|j\rangle\langle j|_A \otimes \Pi_B^{x'} \Delta(\lambda_B)\Pi_B^{x'}) \right]$$
$$= \mathrm{Tr}\left[|j\rangle\langle j|_A \otimes \Delta(\lambda_B) \right] = 1, \tag{B.14}$$

where we have used the fact that $\sum_w \Pi_B^{x'-\frac{w}{\delta w}} = \mathbb{B}_B$ and that both the map E and the map resulting from the application of the projection operators $X \to \sum_{x'} \Pi^{x'} X \Pi^{x'}$ are trace preserving. This proves Eq. (3.25).

We continue with the proof of Condition 2. Exploiting Eq. (3.16), we can expand

$$\Pi^{x'-\frac{w}{\delta w}} = u^{x'-\frac{w}{\delta w}} (u-1)^{n-(x'-\frac{w}{\delta w})} \chi_{x'-\frac{w}{\delta w}}$$
$$= \left\{ u^{x'} (u-1)^{n-x'} \right\} u^{-\frac{w}{\delta w}} (u-1)^{\frac{w}{\delta w}} \chi_{x'-\frac{w}{\delta w}}$$
$$= \left\{ u^{x'} (u-1)^{n-x'} \right\} \left(\frac{u-1}{u} \right)^{\frac{w}{\delta w}} \chi_{x'-\frac{w}{\delta w}}$$
$$= \left\{ u^{x'} (u-1)^{n-x'} \right\} e^{-w} \chi_{x'-\frac{w}{\delta w}} . \tag{B.15}$$

Using the above relation we can write

$$
\begin{aligned}
\sum_{w,j} P(i,w|j)e^w &= \sum_{w,x'} \alpha_{x'}\text{Tr}\big[(|i\rangle\langle i|_A \otimes \chi_{x'-\frac{w}{\delta w}\,B})\mathcal{E}_{AB}(\mathbb{1}_A \otimes \Pi_B^{x'})\big] \\
&\approx \sum_{w,x'} \alpha_{x'-\frac{w}{\delta w}}\text{Tr}\big[(|i\rangle\langle i|_A \otimes \chi_{x'-\frac{w}{\delta w}\,B})\mathcal{E}_{AB}(\mathbb{1}_A \otimes \Pi_B^{x'})\big] \\
&= \text{Tr}\big[(|i\rangle\langle i|_A \otimes \Delta(\lambda_B))\mathcal{E}_{AB}(\mathbb{1}_A \otimes \mathbb{B}_B)\big] \\
&= 1\,,
\end{aligned}
\tag{B.16}
$$

where in the final line we have used the property of the unital map $\mathcal{E}$ to preserve the projector on the diagonal support, while in the second line we have approximated $\alpha_{x'} \approx \alpha_{x'-\frac{w}{\delta w}}$. To verify the validity of the approximation, we invoke Eq. (3.24), which yields

$$
\begin{aligned}
\Big|\sum_{w,j} P(i,w|j)e^w - 1\Big| &\leq \sum_{w,x'} |\alpha_{x'} - \alpha_{x'-\frac{w}{\delta w}}| \\
&\leq \sum_{w:|w|\leq w_{\max}} \sqrt{8\epsilon}\,\frac{|w|}{\delta w} \\
&= \sqrt{8\epsilon}\,f_{\max}(f_{\max}+1)\,,
\end{aligned}
\tag{B.17}
$$

with $f_{\max}$ defined after Eq. (B.3). This shows that Eq. (3.26) is fulfilled in the limit $\epsilon \to 0$.

We conclude by proving Condition 3, that expresses the correspondence between the marginal probabilities and the diagonal components of the states of the system. Using the definition (3.30) of the conditional probability, we can write

$$
\begin{aligned}
\sum_{w,j} P(i,w|j)q_j &= \sum_{j,x'} q_j \alpha_{x'}\text{Tr}\big[(|i\rangle\langle i|_A \otimes \mathbb{B}_B)\mathcal{E}_{AB}(|j\rangle\langle j|_A \otimes \chi_{x'B})\big] \\
&= \text{Tr}\big[(|i\rangle\langle i|_A \otimes \mathbb{B}_B)\mathcal{E}_{AB}(\Delta(\phi_A \otimes \lambda_B))\big] \\
&\approx \text{Tr}\big[(|i\rangle\langle i|_A \otimes \mathbb{B}_B)\mathcal{E}_{AB}(\Delta(\Phi_{AB}))\big] \\
&= \text{Tr}\big[(|i\rangle\langle i|_A \otimes \mathbb{B}_B)\Delta(\Psi_{AB})\big] \\
&= p_i,
\end{aligned}
\tag{B.18}
$$

where we have used Eq. (3.28) and the approximation

$$
\Delta(\Phi_{AB}) \approx \Delta(\phi_A \otimes \lambda_B) = \left(\sum_j q_j|j\rangle\langle j|_B\right) \otimes \left(\sum_{x'} \alpha_{x'}\chi_{x'B}\right).
\tag{B.19}
$$

To verify the validity of the approximation, we invoke Eq. (3.19), which yields

$$
\begin{aligned}
\sum_i \left| \sum_{w,j} P(i,w|j) q_j - p_i \right| &= \sum_i \left| \mathrm{Tr}\big[(|i\rangle\langle i|_A \otimes \mathbb{B}_B) \mathcal{E}_{AB}(\Delta(\phi_A \otimes \lambda_B) - \Delta(\Phi_{AB})) \big] \right| \\
&\leq \frac{1}{2} \| \mathcal{E}_{AB}(\Delta(\phi_A \otimes \lambda_B - \Phi_{AB})) \|_1 \\
&\leq \frac{1}{2} \| \phi_A \otimes \lambda_B - \Phi_{AB} \|_1 \\
&\leq \frac{\epsilon}{2},
\end{aligned}
\tag{B.20}
$$

where we have further used the contractivity of the trace distance under quantum channels and the fact that $\frac{1}{2}\|\rho - \sigma\|_1 = \max_{0 \leq X \leq \mathbb{1}} |\mathrm{Tr}[X(\rho - \sigma)]|$ for two density matrices ρ and σ. This shows that Eq. (3.27) is fulfilled as well in the limit $\epsilon \to 0$.

B.2 Forward and reverse protocols for Crooks' coherence relation

Here we construct and investigate both forward and reverse protocols for battery assisted (S)IO state transformations. The existence of a (S)IO protocol for the state transformation $|\Psi\rangle_{AB} \to |\Phi\rangle_{AB}$, hereby referred to as *forward* protocol, is ensured by the majorisation relation $\Delta(\Psi_{AB}) \prec \Delta(\Phi_{AB})$, which we recall from Eq. (3.4) corresponds to

$$
\Delta(\Psi_{AB}) = \sum_m r_m \Xi_m \Delta(\Phi_{AB}) \Xi_m^\dagger = \mathrm{E}_{AB}(\Delta(\Phi_{AB})).
\tag{B.21}
$$

Let us now define a state $|\Psi'\rangle_{AB}$ with the same diagonal support as $|\Phi\rangle_{AB}$ and a state $|\Phi'\rangle_{AB}$ with the same diagonal support as $|\Psi\rangle_{AB}$, such that the existence of a (S)IO protocol for the state transformation $|\Psi'\rangle_{AB} \to |\Phi'\rangle_{AB}$, hereby referred to as *reverse* protocol, is ensured by the dual map,

$$
\Delta(\Psi'_{AB}) = \sum_m r_m \Xi_m^\dagger \Delta(\Phi'_{AB}) \Xi_m = \mathrm{E}^*_{AB}(\Delta(\Phi'_{AB})),
\tag{B.22}
$$

from which it is explicit that $\Delta(\Psi'_{AB}) \prec \Delta(\Phi'_{AB})$.

According to the B(S)IO framework, as we have seen in Section 3.4, the forward protocol gives rise to a conditional probability distribution $P(i,w|j)$, which takes into account the fact that the coherence in the battery can fluctuate by an amount w with probability $P(w)$. Analogously, we will consider that in the reverse protocol the coherence in the battery is allowed to change by an amount $-w$ with probability $P^{\mathrm{rev}}(-w)$.

Our aim is to compare the coherence distributions in the forward and reverse protocols. To do so, let us adopt the setting of B.1.1, and consider a sequence of forward (S)IO protocols implementing the transformations $|\Psi^{(N)}\rangle_{AB} \to |\Phi^{(N)}\rangle_{AB}$, with initial and final states defined in Eqs. (B.1) and (B.2). In B.1.1 we have introduced a sequence of bistochastic matrices $G^{(N)}$ mapping the (nonzero) diagonal coefficients of $\Phi^{(N)}_{AB}$ to those of $\Psi^{(N)}_{AB}$. We will use the fact that, by construction, these matrices satisfy

$$G^{(N)}(i,z|j,z') = \mathrm{Tr}[(|i\rangle\langle i|_A \otimes |z\rangle\langle z|_B)\mathrm{E}_{AB}(|j\rangle\langle j|_A \otimes |z'\rangle\langle z'|_B)]. \qquad (B.23)$$

Let us now define the transition matrix

$$Q(i,x|j,x') := \mathrm{Tr}[(|i\rangle\langle i|_A \otimes \Pi^x_B)\mathcal{E}_{AB}(|j\rangle\langle j|_A \otimes \chi_{x'B})]. \qquad (B.24)$$

From now on, we will assume for simplicity that $\frac{w}{\delta w}$ is an integer, i.e., $f_w \equiv \frac{w}{\delta w}$. Furthermore, we will impose $x' \in \mathcal{S}$, that is, we will limit the index x' to span within the range in which the diagonal component of the battery has support. In this range, we have specifically

$$
\begin{aligned}
Q(i,x|j,x') &= \sum_{z\in\chi_x, z'\in\chi_{x'}} u^{-x'}(u-1)^{x'-n}\mathrm{Tr}[(|i\rangle\langle i|_A \otimes |z\rangle\langle z|_B)\mathcal{E}_{AB}(|j\rangle\langle j|_A \otimes |z'\rangle\langle z'|_B)] \\
&= \sum_{z\in\chi_x, z'\in\chi_{x'}} G^{(N)}(i,z|j,z') \\
&= \sum_{z\in\chi_x, z'\in\chi_{x'}} u^{-x'}(u-1)^{x'-n}P(i,x|j,x') \\
&= P(i,x|j,x') \\
&= \sum_w P(i,w|j)\delta_{x'-x,\frac{w}{\delta w}}, \qquad (B.25)
\end{aligned}
$$

where in the last line we have used Eq. (B.4). This means that, in the relevant range of x', $Q(i,x|j,x')$ is directly related with the bistochastic matrix mapping the diagonal state coefficients in the forward protocol.

Inspired by Eq. (B.22), we can analogously define a transition matrix for the reverse protocol, given by

$$Q^{\mathrm{rev}}(j,x'|i,x) := \mathrm{Tr}[(|j\rangle\langle j|_A \otimes \Pi^{x'}_B)\mathcal{E}^*_{AB}(|i\rangle\langle i|_A \otimes \chi_{xB})]. \qquad (B.26)$$

Noting that

$$Q^{\mathrm{rev}}(j,x'|i,x) = \frac{u^{x'}(u-1)^{n-x'}}{u^x(u-1)^{n-x}}Q(i,x|j,x'), \qquad (B.27)$$

this can be rewritten as

$$
\begin{aligned}
Q^{\text{rev}}(j, x'|i, x) &= \frac{u^{x'}(u-1)^{n-x'}}{u^x(u-1)^{n-x}} \sum_w P(i, w|j)\delta_{x'-x, \frac{w}{\delta w}} \\
&= \frac{u^{x'}(u-1)^{n-x'}}{u^x(u-1)^{n-x}} \sum_w \delta_{x'-x, \frac{w}{\delta w}} \sum_{x''} \alpha_{x''} \text{Tr}[(|i\rangle\langle i|_A \otimes \Pi_B^{x''-\frac{w}{\delta w}})\mathcal{E}_{AB}(|j\rangle\langle j|_A \otimes \chi_{x''B})] \\
&= \sum_w \delta_{x'-x, \frac{w}{\delta w}} \sum_{x''} \left(\frac{u}{u-1}\right)^{\frac{w}{\delta w}} \alpha_{x''} \text{Tr}[(|i\rangle\langle i|_A \otimes \Pi_B^{x''-\frac{w}{\delta w}})\mathcal{E}_{AB}(|j\rangle\langle j|_A \otimes \chi_{x''B})] \\
&= \sum_w \delta_{x'-x, \frac{w}{\delta w}} \sum_{x''} \alpha_{x''} \text{Tr}[(|j\rangle\langle j|_A \otimes \Pi_B^{x''})\mathcal{E}_{AB}^*(|i\rangle\langle i|_A \otimes \chi_{x''-\frac{w}{\delta w}B})]. \quad \text{(B.28)}
\end{aligned}
$$

This is now approximated to an ideal battery, which gives

$$
\begin{aligned}
Q^{\text{rev}}(j, x'|i, x) &\approx \sum_w \delta_{x'-x, \frac{w}{\delta w}} \sum_{x''} \alpha_{x''+\frac{w}{\delta w}} \text{Tr}[(|j\rangle\langle j|_A \otimes \Pi_B^{x''+\frac{w}{\delta w}})\mathcal{E}_{AB}^*(|i\rangle\langle i|_A \otimes \chi_{x''B})] \\
&\approx \sum_w \delta_{x'-x, \frac{w}{\delta w}} \sum_{x''} \alpha_{x''} \text{Tr}[(|j\rangle\langle j|_A \otimes \Pi_B^{x''+\frac{w}{\delta w}})\mathcal{E}_{AB}^*(|i\rangle\langle i|_A \otimes \chi_{x''B})] \\
&= \sum_w P^{\text{rev}}(j, -w|i)\delta_{x'-x, \frac{w}{\delta w}}, \quad \text{(B.29)}
\end{aligned}
$$

where we have defined the conditional probability distribution for the reverse protocol as

$$
P^{\text{rev}}(j, -w|i) := \sum_{x''} \alpha_{x''} \text{Tr}[(j\rangle\langle j|_A \otimes \Pi_B^{x''+\frac{w}{\delta w}})\mathcal{E}_{AB}^*(|i\rangle\langle i|_A \otimes \chi_{x''B})], \quad \text{(B.30)}
$$

The distribution (B.30) is the counterpart to the distribution $P(i, w|j)$ given in (3.30) for the forward protocol, and it satisfies analogous conditions to those proved in B.1.2, i.e.,

$$
\sum_{j,w} P^{\text{rev}}(j, -w|i) = 1, \quad \text{(B.31)}
$$

$$
\sum_{i,w} P^{\text{rev}}(j, -w|i)e^{-w} = 1, \quad \text{(B.32)}
$$

$$
\sum_{i,w} P^{\text{rev}}(j, -w|i)q_i' = p_j', \quad \text{(B.33)}
$$

where p_i' and q_i' respectively denote the diagonal elements of the initial and final states of the system A in the reverse protocol.

Therefore, given a sequence of forward (S)IO protocols with transition matrix $Q(i, x|j, x')$ mapping the diagonal coefficients of final states $|\Phi^{(N)}\rangle_{AB}$ to those of initial states $|\Psi^{(N)}\rangle_{AB}$, there exists a sequence of reverse (S)IO protocols with transition matrix $Q^{\text{rev}}(j, x'|i, x)$ mapping the diagonal coefficients of final states $|\Phi'^{(N)}\rangle_{AB}$ to those of initial states $|\Psi'^{(N)}\rangle_{AB}$. The initial and final states for the reverse protocol can then be chosen, in analogy to the

analysis of B.1.1, as

$$|\Psi'^{(N)}\rangle_{AB} = \sum_{j=1}^{d}\sum_{x'=0}^{n}\sqrt{\frac{p'_{j,-w}}{N'+1}}\sqsupset_{x'-f_w}(\mathcal{S}')\,|j\rangle_A \otimes |c_{x'}\rangle_B\,, \qquad (B.34)$$

$$|\Phi'^{(N)}\rangle_{AB} = \sum_{i=1}^{d}\sum_{x=0}^{n}\sqrt{\frac{q'_i}{N'+1}}\sqsupset_x(\mathcal{S}')\,|i\rangle_A \otimes |c_x\rangle_B\,, \qquad (B.35)$$

where $p'_{j,-w} := \sum_{i=1}^{d} P^{\mathrm{rev}}(j,-w|i)q'_i$, $\mathcal{S}' := \left\{\frac{n-N'}{2}, ..., \frac{n+N'}{2}\right\}$, and $N' := N - 2f_{\max}$, thus ensuring that $x' \in \mathcal{S}$. It is easy then to verify that Eq. (B.22) is satisfied for the reverse protocol in the limit $N \to \infty$, adapting the derivation of B.1.1.

The coherence analogue of Crooks' theorem can now be obtained. In Crooks' theorem both the initial and final states are thermal, here to derive its equivalent the final states of both the forward and reverse protocols will be taken to be maximally coherent (with diagonal rank d' and d respectively), which sets $q_j = \frac{1}{d'}$ and $q'_i = \frac{1}{d}$.

The forward and reverse coherence distributions are therefore

$$P(w) = \sum_{i,j} P(i,w|j)\frac{1}{d'} \qquad (B.36)$$

$$P^{\mathrm{rev}}(-w) = \sum_{i,j} P^{\mathrm{rev}}(j,-w|i)\frac{1}{d}. \qquad (B.37)$$

Now recalling the definition (3.14) with $x = x' - \frac{w}{\delta w}$, and expanding as in (B.15), we can rewrite Eq. (B.27) simply as

$$Q^{\mathrm{rev}}(j,x'|i,x) = e^w Q(i,x|j,x')\,, \qquad (B.38)$$

from which it is immediate to see that the above two distributions $P(w)$ and $P^{\mathrm{rev}}(-w)$ obey Crooks' analogue relation given in Eq. (3.44).

Appendix C

C.1 Classical treatment

C.1.1 Classical state space and microscopic dynamics

Here, we describe the classical setting with identical particles having an internal spin degree of freedom that is not accessed by the experimenter. The aim is to give a treatment that parallels the quantum one so that the two cases can be compared fairly. Each particle has two degree of freedom – a position $x = 1, \ldots, d$ and a spin $s = 1, \ldots, S$ – which are the accessible and hidden degrees of freedom, respectively. (Note that we only require $S = 2$ in the main text.)

We start from the point of view of a hypothetical observer for whom *all* the particles are fully distinguishable. The effective indistinguishability of the particles will be imposed later by a suitable restriction on the allowed operations. This is rather like the first-quantised description of quantum identical particles. The underlying state space of N distinguishable particles is

$$\Sigma_N = \{(\boldsymbol{x}, \boldsymbol{s}) \mid \boldsymbol{x} \in [d]^N, \boldsymbol{s} \in [S]^N\}, \tag{C.1}$$

where $[k] = \{1, 2, \ldots, k\}$. This can be expressed as a Cartesian product $\Sigma_N = \Sigma_N^x \times \Sigma_N^s$ of the individual spaces for each degree of freedom.

A thermodynamical operation involves coupling the particles to a heat bath and work reservoir, the latter two of which we group into a joint system called the "apparatus" A. This has its own state space Σ_A whose states we designate by a label a. A state of the whole system can therefore be specified by a tuple $(\boldsymbol{x}, \boldsymbol{s}, a)$. We assume the underlying microscopic dynamics to be deterministic and reversible; thus, an evolution of whole

system consists of an invertible mapping

$$(\boldsymbol{x}, \boldsymbol{s}, a) \to (\boldsymbol{x}', \boldsymbol{s}', a'). \tag{C.2}$$

C.1.2   Dynamics independent of spin and particle label

Now we impose the condition that the operation be spin independent. This translates into two features: i) the spins are all unchanged, so $\boldsymbol{s}' = \boldsymbol{s}$, and ii) $\boldsymbol{x}'$ and a' are functions of $\boldsymbol{x}$ and a only, not $\boldsymbol{s}$. It is clear that $\boldsymbol{s}$ is completely decoupled from the other variables, so that the dynamics of the apparatus are the same for any value of $\boldsymbol{s}$. Thus we can drop the redundant information and designate states of the whole system by $(\boldsymbol{x}, a)$.

Next, we impose operational indistinguishability of the particles, again by restricting the allowed operations. An allowed operation must be invariant under a rearrangement of particle labels. For a permutation $\pi \in S_N$, let $\pi[\boldsymbol{x}] = (x_{\pi(1)}, \ldots, x_{\pi(N)})$. Then we require that

$$(\boldsymbol{x}, a) \to (\boldsymbol{x}', a') \Rightarrow (\pi[\boldsymbol{x}], a) \to (\pi[\boldsymbol{x}'], a') \quad \forall \pi \in S_N, \tag{C.3}$$

i.e., the transformation commutes with all permutations. This condition implies that a' is a function only of a and the *type* $\boldsymbol{t}$ of $\boldsymbol{x}$. By this, we mean $\boldsymbol{t} = (t_1, \ldots, t_d)$ specifies the number t_i of particles in each cell i. It is then clear that, as far as the dynamics of A are concerned, it is sufficient to keep track of just $(\boldsymbol{t}, a)$. The total number of effective microstates of the particles, as seen by the ignorant observer, is then the number of possible types, equal to $\binom{N+d-1}{N}$.

C.1.3   Subtlety with overly constrained dynamics

However, there is a subtlety: one can ask whether all (deterministic and reversible) dynamics in the space of $(\boldsymbol{t}, a)$ are possible under the constraint Eq. (C.3). If $(\boldsymbol{t}, a) \to (\boldsymbol{t}', a')$ is possible, then there exist some $\boldsymbol{x}, \boldsymbol{x}'$ of types $\boldsymbol{t}, \boldsymbol{t}'$ respectively such that $(\boldsymbol{x}, a) \to (\boldsymbol{x}', a')$. The condition Eq. (C.3) then determines how all the remaining vectors $\pi[\boldsymbol{x}]$ of type $\boldsymbol{t}$ evolve. There may be a contradiction here – there are two ways in which a transformation might not be possible:

- If there exists π such that $\pi[\boldsymbol{x}] = \boldsymbol{x}$ but $\pi[\boldsymbol{x}'] \neq \boldsymbol{x}'$, then the transformation cannot be deterministic.

- If there exists π such that $\pi[\boldsymbol{x}] \neq \boldsymbol{x}$ but $\pi[\boldsymbol{x}'] = \boldsymbol{x}'$, then the transformation cannot be reversible.

We give the following example, consider $\boldsymbol{x} = (1,1)$, $\boldsymbol{x}' = (1,2)$, which have types $\boldsymbol{t} = (2,0)$, $\boldsymbol{t}' = (1,1)$. A swap of the two particles preserves $\boldsymbol{x}$ but not $\boldsymbol{x}'$ – it is clear that a transition $\boldsymbol{t} \to \boldsymbol{t}'$ cannot be possible. In other words, this is because there is no way of "picking out" a particle from cell 1 and moving it to cell 2 in a way that acts non-preferentially on the particles. In quantum mechanics, this obstacle is avoided because it is possible to act symmetrically on the particles such that the final state is an equal superposition of the two $\boldsymbol{x}' = (1,2)$ and $(2,1)$.

This hints at a way to avoid the problem in the classical case: widening the scope to include stochastic operations. Since it is crucial to require that all dynamics are microscopically deterministic, we introduce stochasticity using additional degrees of freedom containing initial randomness. These couple to the different ways the particles can be permuted, and must necessarily be *hidden*, i.e., not accessible to the observer, in order to maintain ignorance about the particle labels. The idea is to construct globally deterministic, reversible dynamics such that tracing out the hidden degrees of freedom gives stochastic dynamics on $(\boldsymbol{x}, a)$ via the probabilities $p(\boldsymbol{x}', a' | \boldsymbol{x}, a)$. Analogously to Eq. (C.3), we impose the condition

$$p(\boldsymbol{x}', a' | \boldsymbol{x}, a) = p(\pi[\boldsymbol{x}'], a' | \pi[x], a) \quad \forall \pi \in S_N. \tag{C.4}$$

The claim is that such dynamics exist that enable all possible (deterministic, reversible) transformations of $(\boldsymbol{t}, a)$. To see this, consider just one desired transformation $(\boldsymbol{t}, a) \to (\boldsymbol{t}', a')$. We introduce two sets of additional variables $\boldsymbol{h_1}, \boldsymbol{h_2}$ which respectively contain information about $\boldsymbol{x}$ and $\boldsymbol{x}'$. $\boldsymbol{h_1}$ starts in a "ready" state $\boldsymbol{0}$, while $\boldsymbol{h_2}$ is uniformly distributed over all $\boldsymbol{x}'$ of type $\boldsymbol{t}'$. Writing a joint state of all subsystems as $(\boldsymbol{x}, a, \boldsymbol{h_1}, \boldsymbol{h_2})$, it is easily verified that the following dynamics are deterministic and reversible:

$$(\boldsymbol{x}, a, \boldsymbol{0}, \boldsymbol{x}') \to (\boldsymbol{x}', a', \boldsymbol{x}, \boldsymbol{x}') \quad \forall \boldsymbol{x}, \boldsymbol{x}' \text{ of types } \boldsymbol{t}, \boldsymbol{t}', \tag{C.5}$$

where a' is of course a function of $\boldsymbol{t}$ only. Here, $\boldsymbol{h_1}$ keeps a record of the initial configuration (to ensure reversibility) and $\boldsymbol{h_2}$ randomises the final configuration to range uniformly over all $\boldsymbol{x}'$ of type $\boldsymbol{t}'$. Hence we see that $p(\boldsymbol{x}', a' | \boldsymbol{x}, a)$ is constant over all $\boldsymbol{x}, \boldsymbol{x}'$ of interest and thus satisfies condition Eq. (C.4).

Note that $\boldsymbol{h_1}$ has to be initialised in a "pure" state of zero entropy such that it can record information. Such a state, being non-thermal, should be regarded as an additional resource which costs work to prepare. (By contrast, the uniformly random variable $\boldsymbol{h_2}$ is thermal and thus free.) The necessary leakage of information into $\boldsymbol{h_1}$ therefore entails dissipation of work into heat. Hence the work extraction formula (3)[main text] is

technically an upper bound to what can be achieved classically.

This record of information about the initial configuration is seen to be necessary only for those transitions where the set of $\boldsymbol{x}$ of type $\boldsymbol{t}$ is smaller than the set of $\boldsymbol{x}'$ of type $\boldsymbol{t}'$, in order to prevent irreversible merging of states. This situation can be avoided, for instance, in the case of the classical analogue of fermions wherein no more than one particle can occupy a cell. Similarly, in the low density limit (discussion of which appears in the main text), almost all configurations are of this type with very high probability. (One could also argue that this problem is never encountered in reality – as soon as two particles overlap sufficiently, we are already in the quantum parameter regime.)

To summarise what we have shown in this section:

- Classical identical particles can be treated, analogously to the quantum case, as (in principle) distinguishable particles whose dynamics are restricted to be independent of particle label.

- An observer with access only to spin-independent operations can treat the system as if the particles were spin-less.

- There is a subtlety with the particle-label-independent operations that blocks certain transitions. This restriction can be lifted with additional degrees of freedom but may require dissipation of work into heat. This extra cost is zero when particles always occupy distinct cells.

C.2   Details for quantum ignorant observer

In this section, we provide additional details for the entropy change as seen by the ignorant observer.

Recall that Schur-Weyl duality [112, Chapter 5] provides the decomposition

$$\mathcal{H}_x^{\otimes N} = \bigoplus_\lambda \mathcal{H}_x^\lambda \otimes \mathcal{K}_x^\lambda, \qquad (C.6)$$

where λ runs over all Young diagrams containing N boxes and no more than d rows. A Young diagram λ is a set of unlabelled boxes arranged in rows, with non-increasing row length from top to bottom. We can equivalently describe $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_d)$, where λ_i is the number of boxes in row i.   For example, ⊞⊞ would be denoted $(3, 1)$ (where

$N = 4, d = 2$).

$\mathcal{H}_x^\lambda$ and $\mathcal{K}_x^\lambda$ carry irreps of U(d) and S_N respectively, corresponding to irreducible subspaces under the actions of single-particle unitary rotations $u^{\otimes N} \otimes I^{\otimes N}$ and particle label permutations $\Pi \otimes I^{\otimes N}$, each of which act only on the spatial part. The same decomposition works for the spin part $\mathcal{H}_s^{\otimes N}$, although now the Young diagrams λ have maximally two rows. In fact, they correspond to the familiar SU(2) irreps with total angular momentum J, via $\lambda = (N/2 + J, N/2 - J)$.

After putting the spatial and spin decompositions together, projecting onto the overall (anti-)symmetric subspace causes the symmetries of the two components to be linked. For bosons, the overall symmetric subspace (itself a trivial irrep of S_N) occurs exactly once in $\mathcal{K}_x^\lambda \otimes \mathcal{K}_s^{\lambda'}$ if and only if $\lambda = \lambda'$, and otherwise does not [110, Section 7-13]. Thus we have

$$\mathcal{H}_N = \bigoplus_{\lambda, \lambda'} \mathcal{H}_x^\lambda \otimes \mathcal{H}_s^{\lambda'} \otimes P_+ \left[\mathcal{K}_x^\lambda \otimes \mathcal{K}_s^{\lambda'} \right]$$
$$= \bigoplus_\lambda \mathcal{H}_x^\lambda \otimes \mathcal{H}_s^\lambda \quad \text{(bosons).} \tag{C.7}$$

For fermions, the only difference is that the projector P_- onto the antisymmetric subspace enforces $\lambda' = \lambda^T$, denoting the transpose of the Young diagram in which rows and columns are interchanged; thus,

$$\mathcal{H}_N = \bigoplus_\lambda \mathcal{H}_x^{\lambda^T} \otimes \mathcal{H}_s^\lambda \quad \text{(fermions).} \tag{C.8}$$

Due to the use of a two-dimensional spin, we employ the correspondence $J \leftrightarrow \lambda = (N/2 + J, N/2 - J, 0, 0, \dots)$ (with a total of d rows) to replace the label λ by J.

Let us first consider the bosonic case. Thanks to the decomposition in Eq. (C.7), a state ρ (as seen by the informed observer) can be written in terms of the basis $|J, q\rangle_x |J, M\rangle_s |\phi_J\rangle_{xs}$, where $|J, q\rangle \in \mathcal{H}_x^J$, $|J, M\rangle \in \mathcal{H}_s^J$, $|\phi_J\rangle \in \mathcal{K}_x^J \otimes \mathcal{K}_s^J$, as described in the main text. The ignorant observer sees the reduced state after tracing out the spin part, of the form

$$\rho_x = \text{tr}_s \rho = \bigoplus_J p_J \rho_x^J \otimes \text{tr}_s |\phi_J\rangle\langle\phi_J|_{xs}. \tag{C.9}$$

The entropy of this state is

$$S(\rho_x) = H(\boldsymbol{p}) + \sum_J p_J \left[S\left(\rho_x^J\right) + S\left(\text{tr}_s |\phi_J\rangle\langle\phi_J|_{xs}\right) \right], \tag{C.10}$$

where $H(\boldsymbol{p}) := -\sum_J p_J \ln p_J$ is the Shannon entropy of the probability distribution p_J.

As argued in the main text, the fully thermalised final state is of the form

$$\rho'_x = \bigoplus_J p_J \frac{I^J_x}{d_J} \otimes \mathrm{tr}_s |\phi_J\rangle\langle\phi_J|_{xs}, \tag{C.11}$$

with entropy

$$S(\rho'_x) = H(\boldsymbol{p}) + \sum_J p_J \left[\ln d_J + S\left(\mathrm{tr}_s |\phi_J\rangle\langle\phi_J|_{xs}\right)\right]. \tag{C.12}$$

An example of a channel that achieves the mapping from ρ_x to ρ'_x – albeit without a coupling to a heat bath or work reservoir – is the so-called "twirling" operation. This is a probabilistic average over all single-particle unitary rotations $u_x^{\otimes N}$:

$$\mathcal{T}_x(\rho) = \int \mathrm{d}\mu(u_x)\, u_x^{\otimes N} \rho u_x^{\otimes N\dagger}, \tag{C.13}$$

where μ is the Haar measure over the group U(d).

The entropy change for the ignorant observer is therefore

$$\Delta S_{\mathrm{igno}} = S(\rho'_x) - S(\rho_x) = \sum_J p_J \left[\ln d_J - S\left(\rho^J_x\right)\right]. \tag{C.14}$$

(Note that the states ϕ_J do not enter into the entropy change.) Our goal is therefore to determine the probabilities p_J, dimensions d_J, and the entropy of the component states ρ^J_x.

The case of indistinguishable gases is dealt with in the main text: the state is fully in the subspace $J = N/2$, corresponding to the spatially symmetric subspace for bosons and spatially antisymmetric for fermions.

For gases of different spins, the initial state is such that all particles on the left are in $|\uparrow\rangle$ and all on the right are in $|\downarrow\rangle$. Before getting to the thermal state, first consider a pure state in which n_i particles are in each cell i on the left, and m_i in each cell i on the right (such that $\sum_i n_i = n$, $\sum_i m_i = m$). This spatial configuration is denoted by the pair of vectors $(\boldsymbol{n}, \boldsymbol{m})$. The properly symmetrised wavefunction is

$$|\psi(\boldsymbol{n}, \boldsymbol{m})\rangle = \mathcal{N}(\boldsymbol{n}, \boldsymbol{m}) \sum_{\mathrm{distinct}\ \pi \in S_N} \pi|\boldsymbol{n}, \boldsymbol{m}\rangle_x \otimes \pi|\uparrow^n\downarrow^m\rangle_s,$$

$$|\boldsymbol{n}, \boldsymbol{m}\rangle := |1_L^{n_1} 2_L^{n_2} \ldots 1_R^{m_1} 2_R^{m_2} \ldots\rangle, \tag{C.15}$$

where π runs over permutations of the N particles that lead to *distinct* terms $\pi|\boldsymbol{n}, \boldsymbol{m}\rangle_x$. (This is well-defined, since whenever π and π' have the same effect on $|\boldsymbol{n}, \boldsymbol{m}\rangle$, they must also have the same effect on $|\uparrow^n\downarrow^m\rangle$.) $\mathcal{N}(\boldsymbol{n}, \boldsymbol{m})$ is a normalisation factor (such that

$\mathcal{N}(\boldsymbol{n}, \boldsymbol{m})^{-2}$ is the number of distinct terms in the sum). We determine the p_J via the expectation value of the projector P_s^J onto the subspace $\mathcal{H}_s^J$:

$$
\begin{aligned}
&\langle \psi(\boldsymbol{n}, \boldsymbol{m}) | P_s^J | \psi(\boldsymbol{n}, \boldsymbol{m}) \rangle \\
&= \mathcal{N}(\boldsymbol{n}, \boldsymbol{m})^2 \sum_{\text{distinct } \pi, \pi'} \langle \boldsymbol{n}, \boldsymbol{m} | \pi' \pi | \boldsymbol{n}, \boldsymbol{m} \rangle \langle \uparrow^n \downarrow^m | \pi' P_s^J \pi | \uparrow^n \downarrow^m \rangle \\
&= \mathcal{N}(\boldsymbol{n}, \boldsymbol{m})^2 \sum_{\text{distinct } \pi} \langle \uparrow^n \downarrow^m | \pi P_s^J \pi | \uparrow^n \downarrow^m \rangle,
\end{aligned}
\tag{C.16}
$$

where the second line holds because any pair of π, π' giving rise to distinct terms in Eq. (C.15) also have different actions on $|\boldsymbol{n}, \boldsymbol{m}\rangle$. Now we use Clebsch-Gordan coefficients to evaluate each term in this last sum. First note that we can express $|\uparrow^n\rangle$ as a combined spin with $J_1 = M_1 = n/2$, and similarly $|\downarrow^m\rangle$ as a spin with $J_2 = -M_2 = m/2$. The Clebsch-Gordan coefficient $C(\frac{n}{2}, \frac{n}{2}; \frac{m}{2}, \frac{-m}{2}; J, \frac{n-m}{2})$ is precisely the amplitude for this state in the J subspace. This is unchanged by the inclusion of a permutation π, so Eq. (C.16) simplifies to

$$
\langle \psi(\boldsymbol{n}, \boldsymbol{m}) | P_s^J | \psi(\boldsymbol{n}, \boldsymbol{m}) \rangle = \left| C\left(\frac{n}{2}, \frac{n}{2}; \frac{m}{2}, \frac{-m}{2}; J, \frac{n-m}{2} \right) \right|^2.
\tag{C.17}
$$

Now it remains to consider the correct initial state, which is a uniform probabilistic mixture of all $|\psi(\boldsymbol{n}, \boldsymbol{m})\rangle$ with a fixed number of particles n, m on the left and right, respectively. Since the Clebsch-Gordan coefficient is the same for all such configurations, we have [38]

$$
\begin{aligned}
p_J &= \left| C\left(\frac{n}{2}, \frac{n}{2}; \frac{m}{2}, \frac{-m}{2}; J, \frac{n-m}{2} \right) \right|^2 \\
&= \frac{(2J+1) n! m!}{\left(\frac{N}{2} + J + 1 \right)! \left(\frac{N}{2} - J \right)!}.
\end{aligned}
\tag{C.18}
$$

Finally, we determine the entropy of each ρ_x^J component. Using the basis $|J, q\rangle_x |J, M\rangle_s |\phi_J\rangle_{xs}$ provided by the Schur-Weyl decomposition, we have

$$
|\psi(\boldsymbol{n}, \boldsymbol{m})\rangle = \sum_J \sqrt{p_J} |\psi(\boldsymbol{n}, \boldsymbol{m}, J)\rangle_x \left| J, \frac{n-m}{2} \right\rangle_s |\phi_J\rangle_{xs}.
\tag{C.19}
$$

Here, $|\psi(\boldsymbol{n}, \boldsymbol{m}, J)\rangle_x \in \mathcal{H}_x^J$ is some linear combination of the $|J, q\rangle_x$ – without needing to determine these states entirely, it will be sufficient to note that they are orthogonal for different configurations $(\boldsymbol{n}, \boldsymbol{m})$. This follows from the fact that different $|\psi(\boldsymbol{n}, \boldsymbol{m})\rangle$ are fully distinguishable just by measuring the occupation numbers of different cells. Tracing

out s, we find

$$\text{tr}_s \, \psi(\boldsymbol{n}, \boldsymbol{m}) = \bigoplus_J p_J \psi(\boldsymbol{n}, \boldsymbol{m}, J) \otimes \text{tr}_s \, |\phi_J\rangle\langle\phi_J|_{xs},$$
$$\rho_x^J \propto \sum_{\boldsymbol{n}, \boldsymbol{m}} \psi(\boldsymbol{n}, \boldsymbol{m}, J). \tag{C.20}$$

From orthogonality of the $\psi(\boldsymbol{n}, \boldsymbol{m}, J)$, it follows that

$$S(\rho_x^J) = \ln\binom{n + d/2 - 1}{n} + \ln\binom{m + d/2 - 1}{m}. \tag{C.21}$$

Inserted into Eq. (C.14), this results in the claimed entropy changes (9,10)[main text].

C.3   Partial distinguishability

Here, we extend the analysis to include non-orthogonal spins states. As before, we keep the initial spins on the left side of the box as $|\uparrow\rangle^{\otimes n}$, but now on the right we have $|\nearrow\rangle^{\otimes m}$, where $|\nearrow\rangle = \cos(\theta/2) |\uparrow\rangle + \sin(\theta/2) |\downarrow\rangle$.

C.3.1   Informed observer

Let us first discuss the operations allowed to be performed by the informed observer. They are permitted to know about the value of the spins in the $|\uparrow\rangle, |\downarrow\rangle$ basis; they may engineer dynamics diagonal in this basis. Of course, this choice entails a preferred spin basis – this is necessary in order to have a well-defined notion of conditioning dynamics on the value of a spin. We thus require a global unitary of the form $U = \bigoplus_M U_{xsBW}^{(M)}$, where the block structure refers to subspaces with fixed M as defined by the Schur basis. Under a block-diagonal operation, one cannot extract work from coherences between the blocks [127, 160]; that is, the initial state of the spins can be effectively replaced by the dephased state

$$\Phi(\rho_{xs}) := \sum_M Q_s^M \rho_{xs} Q_s^M = \sum_M q_M \rho_{xs}^{(M)}, \tag{C.22}$$

where Q_s^M is the projector onto the M block. In other words, the state behaves thermodynamically as a statistical mixture of the different z-spin numbers M. It follows that the overall entropy change is the average

$$\Delta S_{\text{info}}(\rho_{xs}) = \sum_M q_M \Delta S_{\text{info}}(\rho_{xs}^{(M)}). \tag{C.23}$$

As for the case of orthogonal spins, the initial state is a uniform mixture of states generalising equation Eq. (C.15),

$$|\psi(\boldsymbol{n}, \boldsymbol{m})\rangle = \mathcal{N}(\boldsymbol{n}, \boldsymbol{m}) \sum_{\text{distinct } \pi \in S_N} \pi |\boldsymbol{n}, \boldsymbol{m}\rangle_x \otimes \pi |\uparrow^n \nearrow^m\rangle_s, \tag{C.24}$$

where again it is sufficient (and well-defined) for π to run only over permutations that lead to distinct $\pi |\boldsymbol{n}, \boldsymbol{m}\rangle_x$. As before, $\mathcal{N}^{-2}$ is simply the number of such distinct terms (independent of θ).

Expanding $|\nearrow^m\rangle$ in the preferred basis, it is easily seen that

$$|\nearrow^m\rangle = \sum_{k=0}^{m} \cos(\theta/2)^{m-k} \sin(\theta/2)^k \sum_{\text{distinct } \pi \in S_m} \pi \left|\uparrow^{m-k}\downarrow^k\right\rangle, \tag{C.25}$$

and so

$$q_M = \langle \uparrow^n \nearrow^m | Q_s^M | \uparrow^n \nearrow^m\rangle = \binom{m}{(n+m)/2 - M} \cos(\theta/2)^{m-n+2M} \sin(\theta/2)^{n+m-2M}, \tag{C.26}$$

having used $M = (n+m)/2 - k$. Without needing to know the form of $Q_s^M |\psi(\boldsymbol{n}, \boldsymbol{m})\rangle$, it is sufficient to note that all such states are pure and must be orthogonal, since they can be mutually perfectly distinguished by measuring the occupation number in each cell. The entropy $S(\rho_{xs}^{(M)})$ is therefore just as in Eq. (C.21) for each M.

Due to the block-diagonal structure of the global unitary U, the maximum entropy final state is given by a maximally mixed state for each M block. Considering the number of possible spatial configurations for a fixed number of up and down spins, the dimension of the M block is found to be $\binom{(n+m)/2 - M + d - 1}{(n+m)/2 - M}\binom{(n+m)/2 + m + d - 1}{(n+m)/2 + M}$ in the bosonic case. Hence the overall entropy change is

$$\Delta S_{\text{info}} = \sum_{M=-N/2}^{N/2} q_M \left[\ln \binom{N/2 - M + d - 1}{N/2 - M} + \ln \binom{N/2 + M + d - 1}{N/2 + M} \right] \\ - \left[\ln \binom{n + d/2 - 1}{n} + \ln \binom{m + d/2 - 1}{m} \right].$$ (C.27)

In the fermionic case, analogous counting gives

$$\Delta S_{\text{info}} = \sum_{M=-N/2}^{N/2} q_M \left[\ln \binom{d}{N/2 - M} + \ln \binom{d}{N/2 + M} \right] - \left[\ln \binom{d/2}{n} + \ln \binom{d/2}{m} \right].$$ (C.28)

C.3.2 Ignorant observer

For the ignorant observer, we now have to analyse ρ_x^J. From Eq. (C.24) (recalling that the permutations to be summed over are those that lead to distinct $\pi|\boldsymbol{n}, \boldsymbol{m}\rangle_x$),

$$\text{tr}_s \left[P_s^J |\psi(\boldsymbol{n}, \boldsymbol{m})\rangle\langle\psi(\boldsymbol{n}, \boldsymbol{m})| \right] = \mathcal{N}^2 \sum_{\pi, \pi'} \langle \uparrow^n \nearrow^m | \pi^\dagger P_s^J \pi'| \uparrow^n \nearrow^m \rangle \, \pi' |\boldsymbol{n}, \boldsymbol{m}\rangle\langle\boldsymbol{n}, \boldsymbol{m}| \pi^\dagger$$

$$= \mathcal{N}^2 \sum_{\pi, \pi'} \langle \uparrow^n \nearrow^m | P_s^J \pi^\dagger \pi'| \uparrow^n \nearrow^m \rangle \, \pi' |\boldsymbol{n}, \boldsymbol{m}\rangle\langle\boldsymbol{n}, \boldsymbol{m}| \pi^\dagger,$$ (C.29)

using the fact that the projector P_s^J commutes with permutations. In order to simplify this, we examine coefficients of the form $\langle \uparrow^n \nearrow^m | P_s^J \pi| \uparrow^n \nearrow^m \rangle$. Using the Schur basis just for the spin part, in general one can expand $|\uparrow^n \nearrow^m\rangle = \sum_{J,M,r} \omega_{J,M,r} |J, M, r\rangle$. The r label, representing the part of the basis acted upon by the permutation group, consists of any quantum numbers needed to complete the set along with J and M. We describe a convenient choice of such numbers, denoted $j_1, j_{1,2}, \ldots, j_{1,\ldots,n}$ and $k_1, k_{1,2}, \ldots, k_{1,\ldots,m}$. j_1 is the total spin eigenvalue of spin 1, $j_{1,2}$ of spins 1 and 2 together, and so on. $k_1, \ldots$ have the same meaning, but for the remaining spins label $n + 1, \ldots, n + m$. That these complete the set of quantum numbers is evident from imagining performing an iterated Clebsch-Gordan procedure. This would involve coupling spins 1 and 2, then adding in spin 3, and so on up to spins n. Spins $n + 1$ up to $n + m$ would be coupled recursively in

the same manner, and then finally the two blocks of spins coupled to give the overall J.

For the state $|{\uparrow}^n {\nearrow}^m\rangle$ each of the two blocks of spins is fully symmetric, meaning that each of these spin eigenvalues is maximal: $j_1 = k_1 = \frac{1}{2}$, $j_2 = k_2 = 1$, ..., $j_{1,...,n} = \frac{n}{2}$, $k_{1,...,m} = \frac{m}{2}$. Given this choice of basis, there is only a single value of $r = r_0$ in the expansion of $|{\uparrow}^n {\nearrow}^m\rangle$, referring to this collection of spin eigenvalues. Therefore we can write $|{\uparrow}^n {\nearrow}^m\rangle = \sum_{J,M} \omega_{J,M} |J, M, r_0\rangle$, and

$$
\begin{aligned}
\langle {\uparrow}^n {\nearrow}^m | P_s^J \pi | {\uparrow}^n {\nearrow}^m \rangle &= \sum_{M,M'} \omega^*_{J,M'} \omega_{J,M} \langle J, M', r_0 |\pi| J, M, r_0\rangle \\
&= \sum_M |\omega_{J,M}|^2 \langle J, M, r_0 |\pi| J, M, r_0\rangle \\
&=: \sum_M |\omega_{J,M}|^2 \eta_J(\pi)
\end{aligned}
\tag{C.30}
$$

since $\langle J, M, r_0 |\pi| J, M, r_0\rangle$ is independent of M (and r_0 is fixed anyhow). Expanding $|{\uparrow}^n {\nearrow}^m\rangle$ in the preferred basis and using the Clebsch-Gordan coefficients for coupling the two blocks of spins gives

$$
|\omega_{J,M}|^2 = q_M \left| C\left(\frac{n}{2}, \frac{n}{2}; \frac{m}{2}, M - \frac{n}{2}; J, M\right)\right|^2,
\tag{C.31}
$$

where q_M is defined in Eq. (C.26). Putting this into Eq. (C.29), we have

$$
\begin{aligned}
\mathrm{tr}_s\left[P_s^J |\psi(\boldsymbol{n}, \boldsymbol{m})\rangle\langle\psi(\boldsymbol{n}, \boldsymbol{m})|\right] =& \mathcal{N}^2 \sum_{\pi,\pi'} \sum_M |\omega_{J,M}|^2 \eta_J(\pi^\dagger \pi') \pi' |\boldsymbol{n}, \boldsymbol{m}\rangle\langle\boldsymbol{n}, \boldsymbol{m}| \pi^\dagger \\
=& \sum_M q_M \left| C\left(\frac{n}{2}, \frac{n}{2}; \frac{m}{2}, M - \frac{n}{2}; J, M\right)\right|^2 \\
& \left[\mathcal{N}^2 \sum_{\pi,\pi'} \eta_J(\pi^\dagger \pi') \pi' |\boldsymbol{n}, \boldsymbol{m}\rangle\langle\boldsymbol{n}, \boldsymbol{m}| \pi^\dagger\right].
\end{aligned}
\tag{C.32}
$$

Crucially, the state in brackets is independent of M and must therefore be identical to the state we named $|\psi(\boldsymbol{n}, \boldsymbol{m}, J)\rangle$ in Eq. (C.19). Hence the remaining analysis runs exactly as in the orthogonal spin case, apart from the replacement of p_J by $\sum_M q_M |C\left(\frac{n}{2}, \frac{n}{2}; \frac{m}{2}, M - \frac{n}{2}; J, M\right)|^2$. Thus, all that changes is the probability distribution over J, and this only depends on the probability over M, determined ultimately by the angle θ.

C.4   Dimension counting

From [102, Chapter 7], we have (labelling by λ instead of J)

$$\dim \mathcal{H}_x^\lambda = \frac{\prod_{1\le i<j\le d}(\tilde\lambda_i - \tilde\lambda_j)}{\prod_{m=1}^{d-1} m!},$$
$$\tilde\lambda := \lambda + (d-1, d-2, \ldots, 0). \tag{C.33}$$

First take the bosonic case. Since the Young diagram for the SU(2) spin representation has no more than two rows, the same λ labelling the spatial part has no more than two *non-zero* rows. Hence we have $\tilde\lambda = \left(\frac{N}{2} + J + d - 1, \frac{N}{2} - J + d - 2, d - 3, d - 4, \ldots 0\right)$. Calculating the product in the numerator of Eq. (C.33) is aided by the table below, which lists the values of $\tilde\lambda_i - \tilde\lambda_j$, where i labels the row and $j > i$ labels the column:

	2	3	4	5	...	$d-1$	d
1	$2J+1$	$\frac{N}{2}+J+2$	$\frac{N}{2}+J+3$	...	...	...	$\frac{N}{2}+J+d-1$
2		$\frac{N}{2}-J+1$	$\frac{N}{2}-J+2$	...	...	...	$\frac{N}{2}-J+d-2$
3			1	2	...	...	$d-3$
4				1	...	...	$d-4$
$\vdots$							$\vdots$
$d-2$						1	2
$d-1$							1

$$\tag{C.34}$$

The product of the terms in the first row is

$$(2J+1)\frac{(\frac{N}{2}+J+d-1)!}{(\frac{N}{2}+J+1)!}, \tag{C.35}$$

the second row gives

$$\frac{(\frac{N}{2}-J+d-2)!}{(\frac{N}{2}-J)!}, \tag{C.36}$$

and the remaining rows give

$$\prod_{m=1}^{d-3} m!. \tag{C.37}$$

Putting these into Eq. (C.33) results in the expression for $d_{N,J}^B$ in (11)[main text].

For fermions, we instead use the transpose of the Young diagram, with

$$\lambda^T = (\underbrace{2,\ldots,2}_{\frac{N}{2}-J}, \underbrace{1,\ldots,1}_{2J}). \tag{C.38}$$

An important restriction on λ^T is that the number of rows can never be greater than the dimension, so $\frac{N}{2} + J \leq d$. We find

$$
\tilde{\lambda}^T = \underbrace{(d+1, d, d-1, \ldots, d - \frac{N}{2} + J + 2,}_{\frac{N}{2} - J} \underbrace{d - \frac{N}{2} + J, d - \frac{N}{2} + J - 1, \ldots, d - \frac{N}{2} - J + 1,}_{2J}
$$
$$
\underbrace{d - \frac{N}{2} - J - 1, \ldots, 0)}_{d - \frac{N}{2} - J}. \tag{C.39}
$$

As before, the differences $\tilde{\lambda}^T_i - \tilde{\lambda}^T_j$ can be arranged as follows:

	2	3	...	$\frac{N}{2}-J$	$\frac{N}{2}-J+1$	$\frac{N}{2}-J+2$	...	$\frac{N}{2}+J$	$\frac{N}{2}+J+1$	$\frac{N}{2}+J+2$	...	$d-1$	d
1	1	2	...	$\frac{N}{2}-J-1$	$\frac{N}{2}-J+1$	$\frac{N}{2}-J+2$	...	$\frac{N}{2}+J$	$\frac{N}{2}+J+2$	$\frac{N}{2}+J+3$	...	d	$d+1$
2		1	...	$\frac{N}{2}-J-2$	$\frac{N}{2}-J$	$\frac{N}{2}-J+1$	...	$\frac{N}{2}+J-1$	$\frac{N}{2}+J+1$	$\frac{N}{2}+J+2$	...	$d-1$	d
$\vdots$				$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$
$\frac{N}{2}-J-1$				1	3	4	...	$2J+2$	$2J+4$	$2J+5$	...	$d-(\frac{N}{2}-J)+2$	$d-(\frac{N}{2}-J)+3$
$\frac{N}{2}-J$					2	3	...	$2J+1$	$2J+3$	$2J+4$	...	$d-(\frac{N}{2}-J)+1$	$d-(\frac{N}{2}-J)+2$
$\frac{N}{2}-J+1$						1	...	$2J-1$	$2J+1$	$2J+2$	...	$d-(\frac{N}{2}-J)-1$	$d-(\frac{N}{2}-J)$
$\vdots$								$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$
$\frac{N}{2}+J-1$								1	3	4	...	$d-(\frac{N}{2}+J)+1$	$d-(\frac{N}{2}+J)+2$
$\frac{N}{2}+J$									2	3	...	$d-(\frac{N}{2}+J)$	$d-(\frac{N}{2}+J)+1$
$\frac{N}{2}+J+1$										1	...	$d-(\frac{N}{2}+J)-2$	$d-(\frac{N}{2}+J)-1$
$\vdots$												$\vdots$	$\vdots$
$d-2$												1	2
$d-1$													1

$$\tag{C.40}$$

Here, the blue and red lines indicate the division into the three main index groups. We want to calculate the product of all rows in the table. The bottom group of rows gives

$$
\prod_{m=1}^{d-(N/2+J)-1} m!. \tag{C.41}
$$

The next group up, being careful to discount the terms lost due to the jump at column $j = N/2 + J$, gives

$$
\prod_{m=d-(N/2+J)+1}^{d-(N/2-J)} \frac{m!}{m - (d - (N/2 + J))} \tag{C.42}
$$

Finally, the top group of rows, noting the additional jump at $j = N/2 - J + 1$, gives

$$
\prod_{m=d-(N/2-J)+2}^{d+1} \frac{m!}{[m - (d - (N/2 + J))][m - (d - (N/2 - J) + 1)]}. \tag{C.43}
$$

Inserting into Eq. (C.33), we need to divide the product of the above three terms by $\prod_{m=1}^{d-1} m!$. This factor cancels all the factorials present in the above three expressions,

with the exception of the top two rows, and contributes two factorials occurring at $m = d(N/2 + J)$, $d - (N/2 - J) + 1$. Therefore we have

$$
\begin{aligned}
d_{N,J}^F &= \prod_{r=d-N/2-J+1}^{d-N/2+J} \frac{1}{r - d + N/2 + J} \cdot \prod_{m=d-N/2+J+2}^{d+1} \frac{1}{(m - d + N/2 + J)(m - d + N/2 - J - 1)} \\
&\quad \cdot \frac{d!(d+1)!}{(d - N/2 + J + 1)!(d - N/2 - J)!} \\
&= \frac{1}{(2J)!} \cdot \frac{(2J+1)!}{(N/2 + J + 1)!(N/2 - J)!} \cdot \frac{d!(d+1)!}{(d - N/2 + J + 1)!(d - N/2 - J)!} \\
&= \frac{(2J+1)d!(d+1)!}{(N/2 + J + 1)!(N/2 - J)!(d - N/2 + J + 1)!(d - N/2 - J)!}.
\end{aligned}
\tag{C.44}
$$

C.5   Low density limit

C.5.1   Bosons

Here we prove equation (18)[main text] for bosons. For simplicity, we take $n = m$. The result rests on the observation that, for sufficiently large d, the ratio $d_J^B/p_J \approx \binom{n+d-1}{n}^2$. We have

$$
\begin{aligned}
\frac{d_J^B/p_J}{\binom{n+d-1}{n}^2} &= \frac{(d-1)!(d+n+J-1)!(d+n-J-2)!}{(d-2)!(d+n-1)!^2} \\
&= (d-1)\frac{\prod_{k=0}^{J-1}(d+n+k)}{\prod_{k=0}^{J}(d+n-J-1+k)} \\
&= \left(1 - \frac{1}{d}\right) \prod_{k=0}^{J-1}(1 + [n+k]/d) \prod_{k=0}^{J}(1 + [n-J-1+k]/d)^{-1}
\end{aligned}
\tag{C.45}
$$

Letting $x_k = [n+k]/d$, we have

$$
\begin{aligned}
\prod_{k=0}^{J-1}(1 + [n+k]/d) &= \sum_{k=0}^{J-1} x_k + \sum_{0=k<l}^{J-1} x_k x_l + \mathcal{O}(\epsilon^3) \\
&= \sum_{k=0}^{J-1} x_k + \frac{1}{2}\left[\left(\sum_{k=0}^{J-1} x_k\right)^2 - \sum_{k=0}^{J-1} x_k^2\right] + \mathcal{O}(\epsilon^3) \\
&=: B_1 + B_2 + \mathcal{O}(\epsilon^3),
\end{aligned}
\tag{C.46}
$$

where the first and second order terms are evaluated to be

$$B_1 = \frac{J(2n + J - 1)}{2d}, \tag{C.47}$$

$$B_2 = \frac{J(J-1)(J[12n - 7] + 12n[n-1] + 3J^2 + 2)}{24d^2}, \tag{C.48}$$

and $\epsilon = n^2/d$. Similarly, letting $y_k = [n - J - 1 + k]/d$,

$$\prod_{k=0}^{J}(1 + [n - J - 1 + k]/d) = \sum_{k=0}^{J} y_k + \frac{1}{2}\left[\left(\sum_{k=0}^{J} y_k^2\right)^2 - \sum_{k=0}^{J} y_k^2\right] + \mathcal{O}(\epsilon^3)$$

$$=: C_1 + C_2 + \mathcal{O}(\epsilon^3), \tag{C.49}$$

with

$$C_1 = \frac{(J+1)(2n - J - 2)}{2d}, \tag{C.50}$$

$$C_2 = \frac{J(J+1)(12n^2 - 12n[J+2] + 3J^2 + 11J + 10)}{24d^2}. \tag{C.51}$$

We then have

$$\frac{d_J^B/p_J}{\binom{n+d-1}{n}^2} = \left(1 - \frac{1}{d}\right)(1 + B_1 + B_2)(1 + C_1 + C_2)^{-1} + \mathcal{O}(\epsilon^3)$$

$$= 1 + R_1 + R_2 + \mathcal{O}(\epsilon^3), \tag{C.52}$$

$$R_1 = B_1 - C_1 - \frac{1}{d}$$

$$= \frac{J(J+1) - n}{d}, \tag{C.53}$$

$$R_2 = B_2 - C_2 + C_1^2 - \frac{B_1}{d} + \frac{C_1}{d} - B_1 C_1$$

$$= \frac{2n^2 - 2n(2J[J+1] + 1) + J(J+1)(J^2 + J + 2)}{2d^2}. \tag{C.54}$$

We now use this to compute the deficit in the change of entropy, as compared with the entropy for the informed observer:

$$\Delta S_{\text{igno}} - \Delta S_{\text{info}} = \sum_J p_J \ln\left(\frac{d_x^J/p_J}{\binom{n+d-1}{n}^2}\right) + p_J \ln p_J$$

$$= \sum_J p_J \ln\left(1 + R_1 + R_2 + \mathcal{O}[\epsilon^3]\right) - H(\boldsymbol{p})$$

$$= \sum_J p_J\left(R_1 + R_2 - \frac{R_1^2}{2}\right) + \mathcal{O}(\epsilon^3) - H(\boldsymbol{p}), \tag{C.55}$$

having used the expansion $\ln(1 + x) = x - x^2/2 + \dots$ for small x.

In order to compute the first and second order terms in Eq. (C.55) exactly, we need the following sums involving binomial coefficients:

$$\sum_{J=0}^{n} \binom{2n+1}{n+J+1}(2J+1) = (2n+1)\binom{2n}{n}, \tag{C.56}$$

$$\sum_{J=0}^{n} \binom{2n+1}{n+J+1}(2J+1)J(J+1) = (2n)(2n+1)\binom{2n-1}{n-1}. \tag{C.57}$$

These are both proved using the easily checked identity

$$\frac{N-2k}{N}\binom{N}{k} = \binom{N-1}{k} - \binom{N-1}{k-1}. \tag{C.58}$$

For Eq. (C.56), we have (setting $k = n - J$, $N = 2n+1$)

$$\sum_{J=0}^{n} \binom{2n+1}{n+J+1}(2J+1) = \sum_{J=0}^{n}\binom{2n+1}{n-J}(2J+1)$$

$$= \sum_{k=0}^{n}\binom{2n+1}{k}(2n+1-2k)$$

$$= \sum_{k=0}^{n}(2n+1)\left[\binom{2n}{k} - \binom{2n}{k-1}\right]$$

$$= (2n+1)\binom{2n}{n}. \tag{C.59}$$

Similarly, for Eq. (C.57),

$$\sum_{J=0}^{n} \binom{2n+1}{n+J+1}(2J+1)J(J+1) = \sum_{k=0}^{n}\binom{2n+1}{k}(2n+1-2k)(n-k)(n-k+1)$$

$$= \sum_{k=0}^{n}(2n+1)\left[\binom{2n}{k} - \binom{2n}{k-1}\right](n-k)(n-k+1)$$

$$= (2n+1)\sum_{k=0}^{n}\binom{2n}{k}(n-k)(n-k+1)$$

$$- (2n+1)\sum_{k=0}^{n-1}\binom{2n}{k}(n-k-1)(n-k)$$

$$= (2n+1)\sum_{k=0}^{n-1}\binom{2n}{k}(n-k)\left[(n-k+1) - (n-k-1)\right]$$

$$= (2n+1)\sum_{k=0}^{n-1}\binom{2n}{k}(2n-2k), \tag{C.60}$$

and by using Eq. (C.58) with $N = 2n$,

$$\sum_{J=0}^{n} \binom{2n+1}{n+J+1}(2J+1)J(J+1) = (2n+1)(2n)\sum_{k=0}^{n-1}\left[\binom{2n-1}{k} - \binom{2n-1}{k-1}\right]$$

$$= (2n+1)(2n)\binom{2n-1}{n-1}. \tag{C.61}$$

Recall that

$$p_J = \frac{(n!)^2}{(2n+1)!}\binom{2n+1}{n+J+1}(2J+1), \tag{C.62}$$

so the first order contribution is

$$\begin{aligned}
\sum_{J=0}^{n} p_J R_1(J) &= \sum_{J=0}^{n} p_J \frac{J(J+1) - n}{d} \\
&= -\frac{n}{d} + \frac{(n!)^2}{d(2n+1)!}\binom{2n+1}{n+J+1}(2J+1)J(J+1) \\
&= -\frac{n}{d} + \frac{(n!)^2}{d(2n+1)!}(2n+1)(2n)\binom{2n-1}{n-1} \\
&= -\frac{n}{d} + \frac{(n!)^2(2n+1)(2n)(2n-1)!}{d(2n+1)!(n-1)!(n!)} \\
&= -\frac{n}{d} + \frac{n}{d} = 0. \tag{C.63}
\end{aligned}$$

The second order is

$$\begin{aligned}
\sum_{J} p_J \left[R_2(J) - \frac{R_1(J)^2}{2}\right] &= \sum_{J=0}^{n} p_J \frac{n(n-2) - 2(n-1)J(J+1)}{2d^2} \\
&= \frac{n(n-2)}{2d^2} - \frac{2(n-1)}{2d^2}\sum_{J=0}^{n} p_J J(J+1) \\
&= \frac{n(n-2)}{2d^2} - \frac{2(n-1)}{2d^2}n \\
&= -\frac{n^2}{2d^2}. \tag{C.64}
\end{aligned}$$

Therefore, substituting the above into Eq. (C.55), we have

$$\Delta S_{\text{igno}} - \Delta S_{\text{info}} = -H(\boldsymbol{p}) - \frac{n^2}{2d^2} + \mathcal{O}\left(\frac{n^3}{d^3}\right). \tag{C.65}$$

C.5.2   Fermions

The method is the same as for bosons. We expand $\frac{d_J^F/p_J}{\binom{d}{n}^2}$ to second order. Letting $z_k = [k - n - J]/d$, we have

$$\prod_{k=1}^{J}(1 + [k - n - J]/d) = F_1 + F_2 + \mathcal{O}(\epsilon^3), \tag{C.66}$$

where

$$F_1 = \sum_{k=1}^{J} z_k = \frac{-J(2n + J - 1)}{2d}, \tag{C.67}$$

$$F_2 = \frac{1}{2}\left[F_1^2 - \sum_{k=1}^{J} z_k^2\right] = \frac{J(J-1)(2 + 3J^2 + 12n[n-1] + J[12n - 7])}{24d^2}. \tag{C.68}$$

Similarly, letting $w_k = [k - n + 1]/d$,

$$\prod_{k=0}^{J}(1 + [k - n + 1]/d) = G_1 + G_2 + \mathcal{O}(\epsilon^3), \tag{C.69}$$

where

$$G_1 = \sum_{k=0}^{J} w_k + \frac{(J+1)(J - 2n + 2)}{2d}, \tag{C.70}$$

$$G_2 = \frac{1}{2}\left[G_1^2 - \sum_{k=0}^{J} w_k^2\right] = \frac{J(J+1)(10 + 11J + 3J^2 - 12n[J+2] + 12n^2)}{24d^2}. \tag{C.71}$$

We then have

$$\frac{d_J^F/p_J}{\binom{d}{n}^2} = \left(1 + \frac{1}{d}\right)(1 + F_1 + F_2)(1 + G_1 + G_2)^{-1} + \mathcal{O}(\epsilon^3)$$

$$= 1 + T_1 + T_2 + \mathcal{O}(\epsilon^3), \tag{C.72}$$

$$T_1 = F_1 - G_1 + \frac{1}{d}$$

$$= \frac{-J(J+1) + n}{d}, \tag{C.73}$$

$$T_2 = F_2 - G_2 + G_1^2 + \frac{F_1}{d} - \frac{G_1}{d} - F_1 G_1$$

$$= \frac{2n^2 - 2n(2J[J+1] + 1) + J(J+1)(J^2 + J + 2)}{2d^2}. \tag{C.74}$$

Note that compared with the boson case, $T_1 = -R_1$, $T_2 = R_2$, thus the first order vanishes and we again have

$$\Delta S_{\text{igno}} - \Delta S_{\text{info}} = -H(\boldsymbol{p}) - \frac{n^2}{2d^2} + \mathcal{O}\left(\frac{n^3}{d^3}\right). \tag{C.75}$$

C.6 Entropy $H(\mathbf{p})$ for large particle number

Here, we evaluate the entropy $H(\boldsymbol{p})$ for large particle number. We take $n = m \gg 1$. Starting from Eq. (C.18), we can rewrite

$$\begin{aligned} p_J &= (2J+1)\frac{(n!)^2}{(2n+1)!}\binom{2n+1}{n+J+1} \\ &= (2J+1)\frac{(n!)^2 2^{2n+1}}{(2n+1)!}b(n+J+1), \end{aligned} \tag{C.76}$$

where $b(n+J+1) = 2^{-(2n+1)}\binom{2n+1}{n+J+1}$ follows a binomial distribution with $N+1$ trials and a success probability of $1/2$.

Using Stirling's approximation in the form $n! = \sqrt{2\pi}n^{n+1/2}e^{-n+\mathcal{O}(1/n)}$ [85], we have

$$\begin{aligned} \frac{(n!)^2}{(2n+1)!} &= \frac{n^{2n+1}e^{-2n+\mathcal{O}(1/n)}}{\sqrt{2\pi}(2n+1)^{2n+3/2}e^{-2n-1+\mathcal{O}(1/n)}} \\ &= (\sqrt{2\pi e})\left(\frac{n}{2n+1}\right)^{2n+1}\frac{1}{(2n+1)^{1/2}}[1+\mathcal{O}(1/n)] \\ &= \frac{\sqrt{2\pi e}}{2^{2n+1}\left(1+\frac{1}{2n}\right)^{2n+1}(2n+1)^{1/2}}[1+\mathcal{O}(1/n)] \\ &= \frac{\sqrt{2\pi e}}{2^{2n+1}[e+\mathcal{O}(1/n)](2n+1)^{1/2}}[1+\mathcal{O}(1/n)] \\ &= \frac{1}{2^{2n+1}}\sqrt{\frac{2\pi}{2n+1}}[1+\mathcal{O}(1/n)]. \end{aligned} \tag{C.77}$$

Using a local version of the central limit theorem [189, Chapter VII, Theorem 1], we can approximate $b(n+J+1)$ by a normal distribution with mean $(2n+1)/2$ and variance

$(2n + 1)/4$, obtaining

$$p_J = (2J + 1)\sqrt{\frac{2\pi}{2n + 1}}[1 + \mathcal{O}(1/n)]\left[\frac{e^{-\frac{(J+1/2)^2}{n+1/2}}}{\sqrt{2\pi(2n + 1)/4}} + o(n^{-1/2})\right]$$

$$= (2J + 1)\left[\frac{e^{-\frac{(J+1/2)^2}{n+1/2}}}{n + 1/2} + o(1/n)\right][1 + \mathcal{O}(1/n)]$$

$$= (2J + 1)\frac{e^{-\frac{(J+1/2)^2}{n+1/2}}}{n + 1/2}[1 + o(1)][1 + \mathcal{O}(1/n)]$$

$$= (2J + 1)\frac{e^{-\frac{(J+1/2)^2}{n+1/2}}}{n + 1/2}[1 + o(1)], \tag{C.78}$$

where $o(f)$ denotes an error term going to zero strictly faster than f. Then

$$\ln p_J = \ln(2J + 1) - \ln(n + 1/2) - \frac{(J + 1/2)^2}{n + 1/2} + o(1), \tag{C.79}$$

so the entropy is approximated by

$$H(\boldsymbol{p}) = -\sum_{J=0}^{n} p_J\left[\ln(2J + 1) - \ln(n + 1/2) - \frac{(J + 1/2)^2}{n + 1/2} + o(1)\right]$$

$$= \ln(n + 1/2) + o(1) + [1 + o(1)]\sum_{J=0}^{n}(2J + 1)\frac{e^{-\frac{(J+1/2)^2}{n+1/2}}}{n + 1/2}\left[-\ln(2J + 1) + \frac{(J + 1/2)^2}{n + 1/2}\right]. \tag{C.80}$$

For large n, we expect that the sum can be approximated by an integral. To show this, we can use the simplest version of the Euler-Maclaurin formula:

$$\sum_{J=0}^{n} f(J) = \int_0^n f(x)\,\mathrm{d}x + \int_0^n\left(x - \lfloor x \rfloor - \frac{1}{2}\right)f'(x)\,\mathrm{d}x + \frac{f(0) + f(n)}{2},$$

$$f(x) := (2x + 1)\frac{e^{-\frac{(x+1/2)^2}{n+1/2}}}{n + 1/2}\left[-\ln(2x + 1) + \frac{(x + 1/2)^2}{n + 1/2}\right]. \tag{C.81}$$

Firstly, we have

$$f(0) = \frac{e^{-\frac{1}{4(n+1/2)}}}{n + 1/2} \cdot \frac{1}{4(n + 1/2)} = \mathcal{O}(n^{-2}),$$

$$f(n) = 2e^{-(n+1/2)}\left[-\ln(2n + 1) + (n + 1/2)\right] = \mathcal{O}(ne^{-n}). \tag{C.82}$$

Along these lines, it is not hard to see that shifting the initial point from $x = 0$ to $x = 1/2$ leads to an $o(1)$ error, so we change variables to $y = x + 1/2$ and let $g(y) := f(y - 1/2)$.

Additionally, the upper limit can be extended to infinity with an error which can be verified to be $\mathcal{O}(e^{-n}\text{poly}[n, \ln n])$. For the remainder integral, we let $k = (n + 1/2)^{-1}$ and use

$$
\begin{aligned}
g(y) &= 2ke^{-ky^2}\left[-y\ln(2y) + ky^3\right], \\
g'(y) &= 2ke^{-ky^2}\left[2ky^2\ln(2y) - 2k^2y^4 - \ln(2y) - 1 + 3ky^2\right].
\end{aligned}
\tag{C.83}
$$

Together with $|y - \lfloor y \rfloor - 1/2| \le 1/2$, we have

$$
\begin{aligned}
\left|\int_0^\infty \left(y - \lfloor y \rfloor - \frac{1}{2}\right) g'(y)\,\mathrm{d}y\right| \le{}& \left|\int_0^\infty 2k^3 y \ln(2y) e^{-ky^2}\,\mathrm{d}y\right| \\
&+ \left|\int_0^\infty 2k^3 y^4 e^{-ky^2}\,\mathrm{d}y\right| + \left|\int_0^\infty k\ln(2y)e^{-ky^2}\,\mathrm{d}y\right| \\
&+ \left|\int_0^\infty ke^{-ky^2}\,\mathrm{d}y\right| + \left|\int_0^\infty 3k^2 y^2 e^{-ky^2}\,\mathrm{d}y\right|
\end{aligned}
\tag{C.84}
$$

in which the individual integrals can be evaluated with the highest order being $\mathcal{O}\left(\frac{\ln n}{n}\right) = o(1)$.

Overall, therefore,

$$
\begin{aligned}
\sum_{J=0}^n f(J) &= \int_0^\infty g(y)\,\mathrm{d}y + o(1) \\
&= \frac{1}{2}\left(\ln k + \gamma\right) - \ln 2 + 1 + o(1) \\
&= -\frac{1}{2}\ln n + \frac{\gamma}{2} - \ln 2 + 1 + o(1),
\end{aligned}
\tag{C.85}
$$

where $\gamma = 0.557\ldots$ is the Euler-Mascheroni constant. Putting this into Eq. (C.80),

$$
\begin{aligned}
H(\boldsymbol{p}) &= \frac{1}{2}\ln n + \frac{\gamma}{2} - \ln 2 + 1 + o(1) \\
&= \frac{1}{2}\ln n + 0.595... + o(1).
\end{aligned}
\tag{C.86}
$$

Appendix D

D.1 Form of free states

Here we show that every particle-separable state of N particles is of the first-quantised form

$$\rho^{\bullet} = \sum_i \lambda_i |\psi_i\rangle\langle\psi_i|^{\otimes N}, \ \lambda_i \geq 0. \tag{D.1}$$

By assumption, $\rho^{\bullet}$ is separable, so we can write $\rho^{\bullet} = \sum_i \lambda_i \bigotimes_{k=1}^{N} |\psi_i^k\rangle\langle\psi_i^k|$. Since

$$\bigotimes_{k=1}^{N} |\psi_i^k\rangle \in \operatorname{supp}\rho^{\bullet} \subseteq \mathcal{H}_N, \tag{D.2}$$

each term $\bigotimes_{k=1}^{N} |\psi_i^k\rangle$ is in the symmetric subspace. It follows from this symmetry that all $|\psi_i^k\rangle$ are the same for a given i.

D.2 Appending free states

Theorem. *The operation $\mathcal{E}(\rho) = \rho \otimes \sigma$, which appends a fixed state σ in a new set of m modes, preserves the set of free states if and only if $\sigma = |0\rangle\langle 0|$.*

Proof. It is sufficient to let ρ be the simplest free state, a single particle in a single mode: $\rho = |1\rangle\langle 1|$. $\sigma = \sum_N p_N \sigma^{(N)}$ is arbitrary and may have unbounded particle number. Then

$$\rho \otimes \sigma = \sum_N p_N |1\rangle\langle 1| \otimes \sigma^{(N)}. \tag{D.3}$$

The $(N+1)$-particle component of this state is $|1\rangle\langle 1| \otimes \sigma^{(N)}$. In order to particle-separable,

it must be possible to express as

$$|1\rangle\langle 1| \otimes \sigma^{(N)} = \sum_i \lambda_i\, U_i |N+1, 0, 0 \ldots\rangle\langle N+1, 0, 0 \ldots| U_i^\dagger, \tag{D.4}$$

in terms of some set of $m+1$ modes, with $\lambda_i \geq 0$ and the U_i being free unitaries. The left-hand side has exactly one particle in the first mode and N in the remainder, so the same must be true of every term on the right-hand side. So for each i, $U_i |N+1, 0, \ldots\rangle = |1\rangle |\psi_i\rangle$, which is impossible unless $N = 0$. To see this, note that we can write

$$U_i |N+1, 0 \ldots\rangle \propto (a_1^\dagger + b_i^\dagger)^{N+1} |0\rangle, \tag{D.5}$$

where b_i is some linear combination of annihilation operators on the rightmost N modes. Expanding the bracket $(a_1^\dagger + b_i^\dagger)^{N+1}$, we can never have a single term linear in $a_1^\dagger$ unless $N = 0$.

Therefore $p_N = 0$ for $N \geq 0$, so $\sigma = |0\rangle\langle 0|$. Conversely, it is trivially seen that appending vacuum modes always preserves the set of free states. $\qquad\square$

D.3 Free unitaries

In the following section, we work with states of N particles and always in the first-quantised picture, so we drop the additional notation for convenience.

Theorem. *A unitary U on $\mathcal{H}_N$ maps free states into free states if and only if $U = u^{\otimes N}$.*

Proof. Note that we only specify the restriction of U to $\mathcal{H}_N$ rather than the "full" Hilbert space $\mathcal{H}_1^{\otimes N}$. For example, permutations between particles are not of the given form but have trivial action on the symmetric subspace.

By assumption, for any $|\Psi\rangle = |\psi\rangle^{\otimes N}$, we have $U|\Psi\rangle = |\Phi(\Psi)\rangle := |\phi(\psi)\rangle^{\otimes N}$. Taking an inner product for two arbitrary ψ, ψ':

$$\langle\Psi'|\Psi\rangle = \langle\Phi(\Psi')|\Phi(\Psi)\rangle \Rightarrow \langle\psi'|\psi\rangle^N = \langle\phi(\psi')|\phi(\psi)\rangle^N. \tag{D.6}$$

The Nth root of this gives

$$\langle\phi(\psi')|\phi(\psi)\rangle = \langle\psi'|\psi\rangle\, e^{2\pi i n(\psi, \psi')/N}, \tag{D.7}$$
$$n(\psi, \psi') \in \{0, 1, \ldots, N-1\}.$$

Both sides of this equation must be continuous in ψ, ψ'. But $n(\psi, \psi')$ is a continuous

integer-valued function, so must be constant. In particular, $n(\psi, \psi) = 0$, so we conclude that $n \equiv 0$.

By Wigner's theorem [241], any transformation of states that preserves the inner product must be unitary. Therefore there exists unitary u such that $|\phi(\psi)\rangle = u |\psi\rangle \; \forall \psi$, which proves the result. $\qquad \square$

D.4 Free measurements

As in Appendix D.3, we temporarily drop the first-quantised notation. As a first step in the investigation of non-destructive measurements, we need the following Lemma:

Lemma 1. *Let Π be a projector with support on the symmetric subspace of N particles, i.e. $\Pi = P_N \Pi P_N$, where P_N projects onto $\mathcal{H}_N$. Then Π is non-entangling if and only if there exists a projector π on $\mathcal{H}_1$ such that*

$$\Pi = P_N \pi^{\otimes N} P_N. \tag{D.8}$$

Proof. It is immediate that any Π of the form (D.8) preserves symmetric product states; so we need only prove the converse. We start from the observation that for any $|\psi\rangle \in \mathcal{H}_1$, there is a (normalised) $|\phi\rangle \in \mathcal{H}_1$ such that $\Pi|\psi\rangle^{\otimes N} = c|\phi\rangle^{\otimes N}$, where either $c = 0$ or else $c \neq 0$ and $|\phi\rangle^{\otimes N} \in \operatorname{supp}\Pi$. If $c = 0 \; \forall \, |\psi\rangle$, then $\Pi = 0$ since states of the form $|\psi\rangle^{\otimes N}$ span $\mathcal{H}_N$ [113]. Otherwise, there must exist some $|0\rangle$ such that $|0\rangle^{\otimes N} \in \operatorname{supp}\Pi$.

If $\operatorname{rank}\Pi = 1$, then $\Pi = |0\rangle\langle 0|^{\otimes N}$ and we are done. If $\operatorname{rank}\Pi > 1$, then consider any $|\psi\rangle$ orthogonal to $|0\rangle$. Again, we must have $\Pi|\psi\rangle^{\otimes N} = c|\phi\rangle^{\otimes N}$. Note that

$$
\begin{aligned}
c\langle 0|\phi\rangle^N &= c\langle 0|^{\otimes N}|\phi\rangle^{\otimes N} \\
&= \langle 0|^{\otimes N}\left(\Pi|\psi\rangle^{\otimes N}\right) \\
&= \langle 0|^{\otimes N}|\psi\rangle^{\otimes N} = 0,
\end{aligned}
\tag{D.9}
$$

having used $\Pi|0\rangle^{\otimes N} = |0\rangle^{\otimes N}$. So either $c = 0$, or else $c \neq 0$ and $|\phi\rangle$ is orthogonal to $|0\rangle$. Considering all $|\psi\rangle$ orthogonal to $|0\rangle$, it follows that either $\Pi|\psi\rangle^{\otimes N} = 0$ for all such $|\psi\rangle$, or else there exists $|1\rangle$ orthogonal to $|0\rangle$, with $|1\rangle^{\otimes N} \in \operatorname{supp}\Pi$.

Continuing this procedure, we are able to construct a complete basis $\{|k\rangle\}$ of $\mathcal{H}_1$ such that

$$
|k\rangle^{\otimes N} \in
\begin{cases}
\operatorname{supp}\Pi, & 0 \leq k \leq r-1 \\
\ker\Pi, & r \leq k \leq d-1
\end{cases}
\tag{D.10}
$$

for some r.

Now take an arbitrary $|\psi\rangle \in \mathcal{H}_1$, written in terms of the chosen basis as $|\psi\rangle = \sum_{k=0}^{d-1} \psi_k |k\rangle$. Given the properties of this basis, it follows that

$$\langle k|^{\otimes N} \Pi |\psi\rangle^{\otimes N} = \begin{cases} \langle k|^{\otimes N} |\psi\rangle^{\otimes N} = \psi_k^N, & 0 \le k \le r-1 \\ 0, & r \le k \le d-1. \end{cases} \tag{D.11}$$

But since Π preserves product states, $\Pi|\psi\rangle^{\otimes N} = |\phi\rangle^{\otimes N}$ (where $|\phi\rangle$ need not be normalised). Expressing $|\phi\rangle = \sum_{k=0}^{d-1} \phi_k |k\rangle$, $\langle k|^{\otimes N} |\phi\rangle^{\otimes N} = \phi_k^N$, thus

$$\phi_k = \begin{cases} \psi_k e^{2\pi i n_k/N} & 0 \le k \le r-1 \\ 0, & r \le k \le d-1, \end{cases} \tag{D.12}$$

where $n_k \in \{0, \dots, N-1\}$. In principle, n_k may be a function of $|\psi\rangle$; however, the continuity of the mapping under Π ensures that n_k is continuous and hence constant. Furthermore, since $|\phi\rangle^{\otimes N}$ is invariant under this mapping, we must have $n_k \equiv 0$, so that $\phi_k = \psi_k \, \forall k \le r-1$.

The action of Π on an arbitrary product $|\psi\rangle^{\otimes N}$ is therefore identical to the action of $\pi^{\otimes N}$, where

$$\pi := \sum_{k=0}^{r-1} |k\rangle\langle k|. \tag{D.13}$$

Again, since such product states span $\mathcal{H}_N$, this gives (D.8). $\qquad\square$

Theorem. *Let $\{\Pi_i\}_{i=1}^k$ be a set of non-zero orthogonal projectors onto subspaces of $\mathcal{H}_N$ (where $N > 1$) such that $\sum_{i=1}^k \Pi_i = P_N$ and each Π_i preserves the set of particle-separable states. Then $k = 1$ and*

$$\Pi_1 = P_N. \tag{D.14}$$

Proof. From Lemma 1, there exist projectors π_i such that $\Pi_i = P_N \pi_i^{\otimes N} P_N \, \forall i$. It follows from this that the orthogonality relation $\Pi_i \Pi_j = \delta_{i,j} \Pi_i$ implies $\pi_i \pi_j = \delta_{i,j} \pi_i$. Hence there exist orthogonal $|\psi_i\rangle$ such that $|\psi_i\rangle \in \operatorname{supp} \pi_i$. From these, we construct $|\psi\rangle := \frac{1}{\sqrt{k}} \sum_{i=1}^k |\psi_i\rangle$. The action of Π_i on $|\psi\rangle^{\otimes N}$ is

$$\Pi_i |\psi\rangle^{\otimes N} = (\pi |\psi\rangle)^{\otimes N}, \tag{D.15}$$

from which the completeness relation gives

$$1 = \sum_{i=1}^{k} \langle\psi|^{\otimes N} \Pi_i |\psi\rangle^{\otimes N} = \sum_{i=1}^{k} \langle\psi| \pi_i |\psi\rangle^N. \tag{D.16}$$

Using the form of $|\psi\rangle$, the right-hand side evaluates to

$$\sum_{i=1}^{k} \langle\psi| \pi_i |\psi\rangle^N = \sum_{i=1}^{k} \left(\frac{1}{k}\right)^N = \frac{1}{k^{N-1}}. \tag{D.17}$$

Hence there is a contradiction unless $k = 1$, which forces the single projector to be $\Pi_1 = P_N$. $\qquad\square$

Theorem D.4 says that any non-destructive free projective measurement in the N-particle subspace must be trivial. Extending this to measurements over the whole Fock space, respecting the SSR, shows that only a measurement of the number observable $\hat{N}$ is permissible.

Theorem. *Any destructive measurement respecting the SSR preserves the set of particle-separable states $\mathcal{S}$.*

Proof. It is sufficient to prove this for a single projector. Let the measurement be performed on m_B modes of an $(m_A + m_B)$-mode system, having the action

$$\rho_{AB} \to \sigma_A = \text{Tr}_B \left[(\mathbb{1}_A \otimes \Pi_B)\rho_{AB}\right], \tag{D.18}$$

where Π_B is a projector such that $[\Pi_B, \hat{N}_B] = 0$. Any particle-separable pure state has the form $|\psi\rangle \propto (c^\dagger)^N |0\rangle$, where c is a single-particle annihilation operator. Choosing some orthogonal mode set $\{a_i\}$, where $i = 1, \ldots, m_A$ for the unmeasured modes and $i = m_A + 1, \ldots, m_A + m_B$ for the measured modes, we can write $c = a + b$, where a and b are linear combinations of the unmeasured and measured a_i, respectively. Thus we can effectively treat $|\psi\rangle$ as a two-mode state:

$$\begin{aligned}
|\psi\rangle &= \left(a^\dagger + b^\dagger\right)^N |0\rangle_A |0\rangle_B \\
&= \sum_{N_A} r_{N_A} |N_A\rangle_A |N - N_A\rangle_B,
\end{aligned} \tag{D.19}$$

where the r_{N_A} are coefficients.

Then the post-measurement (unnormalised) state is

$$
\sigma_A = \mathrm{Tr}_B \Big[\sum_{N_A, N_A'} r_{N_A} r_{N_A'}^* (\mathbb{1}_A \otimes \Pi_B) |N_A\rangle\langle N_A'|_A
$$
$$
\otimes |N - N_A\rangle\langle N - N_A'|_B \Big]
$$
$$
= \sum_{N_A, N_A'} r_{N_A} r_{N_A'}^* \langle N - N_A'|_B \Pi_B |N - N_A\rangle_B |N_A\rangle\langle N_A'|_A
$$
$$
= \sum_{N_A, N_A'} r_{N_A} r_{N_A'}^* s_{N_A} \delta_{N_A, N_A'} |N_A\rangle\langle N_A'|_A
$$
$$
= \sum_{N_A} |r_{N_A}|^2 s_{N_A} |N_A\rangle\langle N_A|_A. \tag{D.20}
$$

where we have used the fact that Π_B is diagonal in particle number, $[\Pi_B, \hat{N}_B] = 0$, to give $\langle M|_B \Pi_B |N\rangle_B = s_N \delta_{N,M}$. Hence $\sigma_A \in \mathcal{S}$; the extension to mixed initial states ρ_A follows by linearity. $\square$

D.5 Measures of PE

The following results are used to show that if D satisfies a few straightforward properties, then the resulting measure of PE can be expressed as an average over different particle numbers. We write this in a more abstract form which shows a generalisation to arbitrary resource theories with a block-diagonal structure.

Lemma 2. *Suppose a distance measure D satisfies*

1. *(contractivity) $D(\mathcal{E}(\rho), \mathcal{E}(\sigma)) \leq D(\rho, \sigma)$ under any channel $\mathcal{E}$;*

2. *(joint convexity) $D(\sum_i p_i \rho_i, \sum_i p_i \sigma_i) \leq \sum_i p_i D(\rho_i, \sigma_i)$ for any sets of states ρ_i, σ_i and probabilities p_i;*

3. *(direct sum concavity) $D(\bigoplus_i p_i \rho_i, \bigoplus_i q_i \sigma_i) \geq \sum_i p_i D(\rho_i, \sigma_i)$.*

Then it also satisfies

a. *(direct sum linearity) $D(\bigoplus_i p_i \rho_i, \bigoplus_i p_i \sigma_i) = \sum_i p_i D(\rho_i, \sigma_i)$;*

b. *(ensemble contractivity) $\sum_i p_i D(\rho_i, \sigma_i) \leq D(\rho, \sigma)$, where $\{\mathcal{E}_i\}$ is any quantum instrument, and $\mathcal{E}_i(\rho) = p_i \rho_i$, $\mathcal{E}_i(\sigma) = q_i \sigma_i$.*

Proof. To show (a):

$$
\begin{aligned}
\sum_i p_i D(\rho_i, \sigma_i) &\underset{(3)}{\leq} D\left(\bigoplus_i p_i \rho_i, \bigoplus_i p_i \sigma_i\right) \\
&= D\left(\sum_i p_i \rho_i \otimes |i\rangle\langle i|, \sum_i p_i \sigma_i \otimes |i\rangle\langle i|\right) \\
&\underset{(2)}{\leq} \sum_i p_i D\left(\rho_i \otimes |i\rangle\langle i|, \sigma_i \otimes |i\rangle\langle i|\right) \\
&\underset{(1)}{=} \sum_i p_i D(\rho_i, \sigma_i),
\end{aligned}
$$

where, in the last line, we have used the fact that adding and removing an uncorrelated system are both reversible channels which must therefore leave D unchanged. The left- and right-hand sides are equal, thus the initial inequality must actually be an equality.

To show (b), we construct from the instrument a channel $\mathcal{E}(\rho) = \sum_i \mathcal{E}_i(\rho) \otimes |i\rangle\langle i|$, so that

$$
\begin{aligned}
\sum_i p_i D(\rho_i, \sigma_i) &\underset{(3)}{\leq} D\left(\bigoplus_i p_i \rho_i, \bigoplus_i q_i \sigma_i\right) \\
&= D\left(\sum_i p_i \rho_i \otimes |i\rangle\langle i|, \sum_i q_i \sigma_i \otimes |i\rangle\langle i|\right) \\
&= D\left(\mathcal{E}(\rho), \mathcal{E}(\sigma)\right) \\
&\underset{(1)}{\leq} D(\rho, \sigma).
\end{aligned}
$$

$\square$

From this, we obtain:

Theorem. *Suppose that D satisfies properties (1,2,3) listed in Lemma 2. Let F be any convex set of states, and define*

$$
M^D(\rho) := \min_{\sigma \in F} D(\rho, \sigma). \tag{D.21}
$$

Then M^D is an ensemble monotone under instruments $\{\mathcal{E}_i\}$ such that each $\mathcal{E}_i$ preserves the set F.

Furthermore, if $F = \bigoplus_N F_N$, where each F_N is a convex set of states, then

$$
M^D\left(\bigoplus_N p_N \rho^{(N)}\right) = \sum_N p_N M_N^D(\rho^{(N)}), \tag{D.22}
$$

where M_N^D is defined similarly to M^D, but minimising over states in F_N.

Proof. For the first part, we take τ to be the closest state to ρ in F. For any instrument $\{\mathcal{E}_i\}$, let $p_i\rho_i = \mathcal{E}_i(\rho)$, $q_i\tau_i = \mathcal{E}_i(\tau)$. Then

$$
\begin{aligned}
M^D(\rho) &= D(\rho, \tau) \\
&\underset{(b)}{\geq} \sum_i p_i D(\rho_i, \tau_i) \\
&\geq \sum_i p_i \min_{\sigma_i \in F} D(\rho_i, \sigma_i) \\
&= \sum_i p_i M^D(\rho_i).
\end{aligned}
$$

For the second part,

$$
\begin{aligned}
M^D &\left(\bigoplus_N p_N \rho^{(N)} \right) \\
&= \min_{\{q_N, \sigma^{(N)} \in F_N\}} D\left(\bigoplus_N p_N \rho^{(N)}, \bigoplus_N q_N \sigma^{(N)} \right) \\
&\underset{(3)}{\geq} \sum_N p_N \min_{\sigma^{(N)} \in F_N} D\left(\rho^{(N)}, \sigma^{(N)} \right) \\
&\underset{(2)}{\geq} \min_{\{\sigma^{(N)} \in F_N\}} D\left(\bigoplus_N p_N \rho^{(N)}, \bigoplus_N p_N \sigma^{(N)} \right),
\end{aligned}
$$

which shows that the closest state can be chosen to have $q_N = p_N$. Finally, we use (a). $\square$

The relative entropy $S(\rho||\sigma) := \mathrm{Tr}[\rho \log \rho - \rho \log \sigma]$ satisfies all three assumptions of Lemma 2 – in particular, (3) follows from

$$
S\left(\bigoplus_i p_i \rho_i || \bigoplus_i q_i \sigma_i \right) = \sum_i p_i S(\rho_i || \sigma_i) + H(\{p_i\}||\{q_i\}), \tag{D.23}
$$

where the last term is the classical relative entropy (or Kullback-Leibler divergence). Hence the relative entropy measure of PE is

$$
M_{\mathrm{PE}}^{RE}(\rho) = \sum_N p_N M_{\mathrm{PE}}^{RE}(\rho^{(N)}). \tag{D.24}
$$

The same property also holds for distances defined by Schatten p-norms, $D_p(\rho, \sigma) = \|\rho - \sigma\|_p$ [57].

D.6   Monotonicity of metrological measure

The proof of monotonicity of $M_{\mathrm{PE}}^{\mathcal{F}}$ makes use of the following Lemma (which is to our knowledge novel):

Lemma 3. *Let Π be a projector such that $\Pi\rho = \rho$. Then*

$$\mathcal{F}(\rho, H) = \mathcal{F}(\rho, \Pi H \Pi) + 4V(\rho, H) - 4V(\rho, \Pi H \Pi). \qquad \text{(D.25)}$$

Proof. Given the spectral decomposition $\rho = \sum_{i=0}^{d-1} \lambda_i |i\rangle\langle i|$, we have $\lambda_i \Pi |i\rangle = \Pi\rho |i\rangle = \lambda_i |i\rangle$, so $\Pi |i\rangle = |i\rangle \ \forall |i\rangle \in \operatorname{supp}\rho$. Therefore we can write $\Pi = \sum_{i<r} |i\rangle\langle i|$, such that $\lambda_j = 0 \ \forall j \geq r$, where $r = \operatorname{rank}\Pi \geq \operatorname{rank}\rho$. It follows that

$$
\begin{aligned}
\mathcal{F}(\rho, H) &= 2\sum_{i,j} \frac{(\lambda_i - \lambda_j)^2}{\lambda_i + \lambda_j} |\langle i| H |j\rangle|^2 \\
&= 2\sum_{i,j<r} \frac{(\lambda_i - \lambda_j)^2}{\lambda_i + \lambda_j} |\langle i| H |j\rangle|^2 \\
&\quad + 4\sum_{i<r, j\geq r} \frac{(\lambda_i - 0)^2}{\lambda_i + 0} |\langle i| H |j\rangle|^2 \\
&= 2\sum_{i,j<r} \frac{(\lambda_i - \lambda_j)^2}{\lambda_i + \lambda_j} |\langle i| \Pi H \Pi |j\rangle|^2 \\
&\quad + 4\sum_{i<r, j\geq r} \frac{(\lambda_i - 0)^2}{\lambda_i + 0} |\langle i| H |j\rangle|^2 \\
&= \mathcal{F}(\rho, \Pi H \Pi) + 4\sum_{i<r, j\geq r} \lambda_i \langle i| H |j\rangle \langle j| H |i\rangle \\
&= \mathcal{F}(\rho, \Pi H \Pi) + 4\sum_{i<r} \lambda_i \langle i| H \left(\sum_{j\geq r} |j\rangle\langle j| \right) H |i\rangle \\
&= \mathcal{F}(\rho, \Pi H \Pi) + 4\sum_{i<r} \lambda_i \langle i| H(I - \Pi) H |i\rangle \\
&= \mathcal{F}(\rho, \Pi H \Pi) + 4\operatorname{Tr}(\rho H^2) - 4\operatorname{Tr}(\rho H \Pi H) \\
&= \mathcal{F}(\rho, \Pi H \Pi) + 4\operatorname{Tr}(\rho H^2) - 4\operatorname{Tr}(\rho[\Pi H \Pi]^2) \\
&= \mathcal{F}(\rho, \Pi H \Pi) + 4V(\rho, H) - 4V(\rho, \Pi H \Pi), \qquad \text{(D.26)}
\end{aligned}
$$

where the last line uses $\operatorname{Tr}(\rho \Pi H \Pi) = \operatorname{Tr}(\rho H)$. $\qquad \square$

Theorem 4 (main text). *$M_{PE}^{\mathcal{F}}$ is convex and satisfies $M_{PE}^{\mathcal{F}}(\rho) = 0 \ \forall \rho \in \mathcal{S}$. Moreover, let $\mathcal{E}_{S\to SM} \in \mathcal{O}$ contain a single measurement round, such that no conditional operations are performed after the measurement. We may write $\mathcal{E}_{S\to SM}(\rho_S) = \sum_m \mathcal{E}^m(\rho_S) \otimes |m\rangle\langle m|_M$,*

where $\mathcal{E}^m$ is the operation applied to ρ_S conditioned on outcome m. Then

$$M_{PE}^{\mathcal{F}}(\rho_S) \geq M_{PE}^{\mathcal{F}}(\mathcal{E}_{S \to SM}[\rho]).$$ (D.27)

Proof. Convexity of $M_{\mathrm{PE}}^{\mathcal{F}}$ follows from convexity of both the QFI and the function $[\cdot]^+$, and concavity of the variance:

$$
\begin{aligned}
M_{\mathrm{PE}}^{\mathcal{F}}(p\rho + (1-p)\sigma) &\leq \max_h \left[p\mathcal{F}(\rho, H) + (1-p)\mathcal{F}(\sigma, H) \right. \\
&\qquad \left. -4pV(\rho, h) - 4(1-p)V(\sigma, h) \right]^+ \\
&\leq \max_h p\left[\mathcal{F}(\rho, H) - 4V(\rho, h)\right]^+ \\
&\qquad + (1-p)\left[\mathcal{F}(\sigma, H) - 4V(\sigma, h)\right]^+ \\
&\leq pM_{\mathrm{PE}}^{\mathcal{F}}(\rho) + (1-p)M_{\mathrm{PE}}^{\mathcal{F}}(\sigma).
\end{aligned}
$$ (D.28)

We break the proof of monotonicity into the three stages of a particle-separable operation without feed-forward: i) appending modes in the vacuum state; ii) performing a global passive linear unitary; iii) destructively measuring a set of modes.

i) Appending modes in the vacuum state: We append to the system modes S a set of vacuum ancilla modes A. Our aim is to show that

$$M_{\mathrm{PE}}^{\mathcal{F}}(\rho_S \otimes |0\rangle\langle0|_A) = M_{\mathrm{PE}}^{\mathcal{F}}(\rho_S).$$ (D.29)

The proof consists of showing that the optimal observable for the vacuum-added state always acts solely on S. Note that the single-particle Hilbert space of SA splits into $\mathcal{H}_1 = \mathcal{H}_{1,S} \oplus \mathcal{H}_{1,A}$; we denote the projectors onto these subspaces by Π_S, Π_A respectively. Thus any h can be decomposed into the terms

$$h = \Pi_S h \Pi_S + \Pi_A h \Pi_A + \Pi_S h \Pi_A + \Pi_A h \Pi_S =: h' + g' + f + f^\dagger.$$ (D.30)

Each term gives rise to its own second-quantised observable exactly as for H, i.e. $H'^{(N)\bullet} = \sum_{i=1}^N h'_i/\sqrt{N}$ and so on.

We apply Lemma 3 using H and the projector $\Pi = I_S \otimes |0\rangle\langle0|_A$. It may be seen that in first quantisation, $\Pi^{(N)\bullet} = \Pi_S^{\otimes N}$, so that each particle is projected on the subspace $\mathcal{H}_{1,S}$.

Therefore we see that $\Pi H \Pi = H'$. Thus

$$
\begin{aligned}
\mathcal{F}(\rho_S \otimes |0\rangle\langle 0|_A, H) &= \mathcal{F}(\rho_S, H') + 4V(\rho_S \otimes |0\rangle\langle 0|_A, H) \\
&\quad - 4V(\rho_S, H') \\
&= \mathcal{F}(\rho_S, H') + 4\operatorname{Tr}\big(\rho_S \otimes |0\rangle\langle 0|_A H^2\big) \\
&\quad - 4\operatorname{Tr}\big(\rho_S H^2\big) \\
&= \mathcal{F}(\rho_S, H') \\
&\quad + 4\operatorname{Tr}\big(\rho_S \otimes |0\rangle\langle 0|_A [\Pi H^2 \Pi - H'^2]\big)
\end{aligned}
\tag{D.31}
$$

using $\operatorname{Tr}(\rho_S \otimes |0\rangle\langle 0|_A H) = \operatorname{Tr}(\rho_S H')$ for the second line. Now one can also see that $\Pi H^2 \Pi = H'^2 + \Pi F F^\dagger \Pi$, so

$$
\mathcal{F}(\rho_S \otimes |0\rangle\langle 0|_A, H) = \mathcal{F}(\rho_S, H') + 4\operatorname{Tr}\big(\rho_S \otimes |0\rangle\langle 0|_A \Pi F F^\dagger \Pi\big).
\tag{D.32}
$$

From $(FF^\dagger)^{(N)\bullet} = \frac{1}{N}\sum_{i,j=1}^N f_i f_j^\dagger$, it follows that

$$
\Pi_S^{\otimes N}(FF^\dagger)^{(N)\bullet}\Pi_S^{\otimes N} = \frac{1}{N}\sum_{i=1}^N f_i f_i^\dagger,
\tag{D.33}
$$

since $\Pi_S f \Pi_S = 0$ but $\Pi_S f f^\dagger \Pi_S \neq 0$. Consequently,

$$
\operatorname{Tr}\big(\rho_S \otimes |0\rangle\langle 0|_A \Pi F F^\dagger \Pi\big) = \langle f f^\dagger \rangle_{\rho_S \otimes |0\rangle\langle 0|_A}.
\tag{D.34}
$$

Next we have

$$
\begin{aligned}
\mathcal{F}(\rho_S &\otimes |0\rangle\langle 0|_A, H) - 4V(\rho_S \otimes |0\rangle\langle 0|_A, h_1) \\
&= \mathcal{F}(\rho_S, H') + 4\langle ff^\dagger - h^2 \rangle_{\rho_S \otimes |0\rangle\langle 0|_A} + 4\langle h \rangle^2_{\rho_S \otimes |0\rangle\langle 0|_A} \\
&= \mathcal{F}(\rho_S, H') - 4\langle h'^2 \rangle_{\rho_S} + 4\langle h' \rangle^2_{\rho_S} \\
&= \mathcal{F}(\rho_S, H') - 4V(\rho_S, h').
\end{aligned}
\tag{D.35}
$$

Now $\|h'\| = \|\Pi_S h \Pi_S\| \le \|h\|\|\Pi_S\| = \|h\|$. If $\|h'\| = 0$, then both sides of (D.35) are zero and there is nothing left to prove; otherwise, we define $\tilde{h} := h'/\|h'\|$, which has unit norm. Putting this into (D.35) gives

$$
\begin{aligned}
\mathcal{F}(\rho_S &\otimes |0\rangle\langle 0|_A, H) - 4V(\rho_S \otimes |0\rangle\langle 0|_A, h) \\
&= \|h'\|^2 \left[\mathcal{F}(\rho_S, \tilde{H}) - 4V(\rho_S, \tilde{h}) \right] \\
&\le \left[\mathcal{F}(\rho_S, \tilde{H}) - 4V(\rho_S, \tilde{h}) \right]^+ \\
&\le M_{\mathrm{PE}}^{\mathcal{F}}(\rho_S).
\end{aligned}
\tag{D.36}
$$

Maximising over h gives $M_{\mathrm{PE}}^{\mathcal{F}}(\rho_S \otimes |0\rangle\langle 0|_A) \leq M_{\mathrm{PE}}^{\mathcal{F}}(\rho_S)$. Conversely, it is clear that equality is obtained by taking for $\rho_S \otimes |0\rangle\langle 0|_A$ the same observable that maximises the quantity for ρ_S. Thus we have established (D.29).

ii) Passive linear unitaries: $M_{\mathrm{PE}}^{\mathcal{F}}$ is explicitly invariant under such unitaries, since these correspond to a rotation of the single-particle basis, and thus just a basis change for h.

iii) Destructive measurement: We start with a state ρ_{SA} on two sets of modes S, A, where the latter ancilla modes are to be measured with a complete POVM $\{E_m\}_M$ respecting the particle-number SSR. The measurement is represented with a quantum-classical channel taking A to a classical memory M:

$$\rho'_{SM} := \mathcal{E}_{A \to M}(\rho_{SA}) := \sum_m \mathrm{Tr}_A[E_{m,A}\rho_{SA}] \otimes |m\rangle\langle m|_M. \qquad (D.37)$$

For any given h acting only on S, we have

$$M_{\mathrm{PE}}^{\mathcal{F}}(\rho_{SA}) \geq [\mathcal{F}(\rho_{SA}, H) - 4V(\rho_{SA}, h)]^+$$
$$\geq [\mathcal{F}(\rho'_{SM}, H) - 4V(\rho_{SA}, h)]^+. \qquad (D.38)$$

The second inequality follows from the property of $\mathcal{F}(\rho, H)$ being monotonically non-increasing under operations covariant with respect to the observable H [250]. Here, co-variance holds because $\mathcal{E}_{A \to M}$ acts on a different subsystem from H. Next, we see that the variance part is unchanged since the statistics of h do not depend on operations performed on subsystem A, so

$$M_{\mathrm{PE}}^{\mathcal{F}}(\rho_{SA}) \geq [\mathcal{F}(\rho'_{SM}) - 4V(\rho'_{SM}, h)]^+. \qquad (D.39)$$

Finally, maximising the right-hand side over all h gives $M_{\mathrm{PE}}^{\mathcal{F}}(\rho_{SA}) \geq M_{\mathrm{PE}}^{\mathcal{F}}(\rho'_{SM})$ $\qquad \square$

D.7 SSR-entanglement

The activation protocol converts particle entanglement into entanglement that is of use to two parties A, B who are limited to local *covariant* operations that respect the SSR and classical communication.

Definition 1. *[205, 204] An operation between two or more parties is said to be* covariant-LOCC *when it is composed of local operations respecting the local superselection rule, and classical communication.*

Although not spelled out explicitly by [205, 204], the free states of this resource theory

(in a bipartite setting; easily generalised) are the following:

Definition 2. *A bipartite state ρ_{AB} is free in the resource theory of SSR-entanglement when it can be written in the form*

$$\rho_{AB} = \sum_i p_i \rho_A^i \otimes \rho_B^i \tag{D.40}$$

such that each ρ_A^i, ρ_B^i respects the SSR, i.e., $\Phi_S(\rho_S^i) = \rho_S^i$, $S = A, B$. Such a free state is said to be invariant-separable *(since it is invariant under local phase rotations).*

Of course every invariant-separable state is separable, but not vice-versa. This set of free states may be motivated as being those accessible from a given primitive state, such as the vacuum $|0\rangle |0\rangle$ under covariant-LOCC.

Lemma 4. *The following statements are equivalent:*

1. *ρ_{AB} is invariant-separable.*

2. *$\rho_{AB} = \sum_i p_i \psi_A^i \otimes \psi_B^i$ where each ψ_A^i, ψ_B^i is pure and contains a definite number of particles.*

3. *ρ_{AB} is separable and satisfies the local SSR constraint $(\Phi_A \otimes \Phi_B)(\rho_{AB}) = \rho_{AB}$.*

4. *$(\Phi_A \otimes \Phi_B)(\rho_{AB}) = \rho_{AB}$ and, for each N_A, N_B, the local-number projected state $(P_{N_A} \otimes P_{N_B})\rho_{AB}(P_{N_A} \otimes P_{N_B})$ is separable.*

Proof. The equivalence of (1) and (2) is easily seen from the fact that every local-SSR-respecting state $\rho_A^i = \Phi_A(\rho_A^i)$ can be written as a mixture of pure states of definite number. (1) $\Rightarrow$ (3) is also straightforward. Conversely, suppose (3) holds, then we have $\rho_{AB} = \sum_i p_i \rho_A^i \otimes \rho_B^i$ for arbitrary states ρ_A^i, ρ_B^i. But then the local SSR constraint implies that $\rho_{AB} = \sum_i p_i \sigma_A^i \otimes \sigma_B^i$, where $\sigma_S^i = \Phi_S(\rho_S^i)$. Thus (3) $\Rightarrow$ (1).

It is clear that (4) $\Rightarrow$ (3), since

$$(\Phi_A \otimes \Phi_B)(\rho_{AB}) = \sum_{N_A, N_B} (P_{N_A} \otimes P_{N_B})\rho_{AB}(P_{N_A} \otimes P_{N_B}), \tag{D.41}$$

so that if each term in the RHS is separable, then the LHS also is.

Finally, we show that (1) $\Rightarrow$ (4). We have

$$(P_{N_A} \otimes P_{N_B})\rho_{AB}(P_{N_A} \otimes P_{N_B})$$
$$= \sum_i p_i \left(P_{N_A}\rho_A^i P_{N_A}\right) \otimes \left(P_{N_B}\rho_B^i P_{N_B}\right), \tag{D.42}$$

which is separable. □

A state can fail to be invariant-separable in two different (but not mutually exclusive) ways: it may break the local SSR, or it may be entangled. The measures of SSR-entanglement defined here capture the amount of entanglement accessible from a single copy of the state under the local SSR. However, there are states which have $E_{\text{SSR}} = 0$ yet are not invariant-separable – for example, product states which break the local SSR.

Lemma 5. *The distance-based measure of SSR-entanglement can be calculated by a restricted optimisation over SSR-separable states:*

$$E_{SSR}^{D}(\rho) = \min_{\sigma \in \ inv.\text{-}sep.} D(\Phi_A \otimes \Phi_B[\rho_{AB}], \sigma_{AB}). \tag{D.43}$$

Equivalently, the closest separable state to $(\Phi_A \otimes \Phi_B)(\rho_{AB})$ *is invariant-separable.*

Proof. Let $E_{\text{SSR}}^{\prime D}$ be the quantity defined by the right-hand side of (D.43). We prove an inequality in both directions. Since invariant-separable states form a subset of separable states, it is clear that $E_{\text{SSR}}^{\prime D} \geq E_{\text{SSR}}^{D}$. Conversely,

$$
\begin{aligned}
E_{\text{SSR}}^{D}(\rho_{AB}) &= \min_{\sigma \in \ \text{sep.}} D(\Phi_A \otimes \Phi_B[\rho_{AB}], \sigma_{AB}) \\
&\geq \min_{\sigma \in \ \text{sep.}} D(\Phi_A \otimes \Phi_B[\rho_{AB}], \Phi_A \otimes \Phi_B[\sigma_{AB}]) \\
&\geq \min_{\tau \in \ \text{inv.-sep.}} D(\Phi_A \otimes \Phi_B[\rho_{AB}], \tau_{AB}) \\
&= E_{\text{SSR}}^{\prime D}(\rho_{AB}),
\end{aligned}
\tag{D.44}
$$

where we have used the monotonicity of D under $\Phi_A \otimes \Phi_B$ and the fact that $\Phi_A \otimes \Phi_B(\sigma_{AB})$ is invariant-separable. □

A useful consequence of Theorem D.5 is that the relative entropy measure of SSR-entanglement can be written as

$$
\begin{aligned}
E_{\text{SSR}}^{\text{RE}}&(\rho_{AB}) \\
&= \sum_{N_A, N_B} p_{N_A, N_B} E_{\text{SSR}}^{\text{RE}} \left(\frac{(P_{N_A} \otimes P_{N_B})\rho_{AB}(P_{N_A} \otimes P_{N_B})}{p_{N_A, N_B}} \right) \\
&= \sum_{N_A, N_B} p_{N_A, N_B} E^{\text{RE}} \left(\frac{(P_{N_A} \otimes P_{N_B})\rho_{AB}(P_{N_A} \otimes P_{N_B})}{p_{N_A, N_B}} \right),
\end{aligned}
\tag{D.45}
$$

where $p_{N_A, N_B} = \text{Tr}[(P_{N_A} \otimes P_{N_B})\rho_{AB}]$. This measure is seen to provide an extension of the pure-state measure defined by Wiseman and Vaccaro [243].

D.8 Activation protocol

The following Lemma shows that a unitary activation operation can be expressed in a simplified form.

Lemma 6. *Let an activation operation $\mathcal{E}_{C \to AB} \in \mathcal{O}$ map its input m modes on C directly onto A, attach the same number m of vacuum modes in B and interact the two sets by a passive linear unitary U:*

$$\sigma_{AB} = \mathcal{E}_{C \to AB}(\rho_A) = U(\rho_A \otimes |0\rangle\langle 0|_B)U^\dagger. \qquad (D.46)$$

Up to local free unitaries, σ_{AB} is equivalent to the state obtained by replacing U with DV_A, where V_A is a free unitary on the A modes and D is a set of beam splitters acting in parallel, with the action

$$D^\dagger a_i D = r_i a_i + t_i b_i, \quad r_i = \sqrt{1 - t_i^2} \in [0,1], \ i = 1, \ldots, m. \qquad (D.47)$$

Proof. Lemma 2 of [247] shows that U can be decomposed as $W_A W_B D V_A V_B$, where $V_{A,B}$, $W_{A,B}$ are free unitaries acting locally on their respective subsystems. Up to final local unitaries, we can replace this by $D V_A V_B$; moreover, V_B can be removed since it leaves the initial vacuum state $|0\rangle_B$ unchanged. $\square$

It is worth noting that the number of vacuum modes introduced can always be assumed to be no greater than m – again, as a consequence of Lemma 2 in [247].

The faithfulness of the activation is proven below for almost all such unitaries (apart from those with vanishing beam-splitter parameters).

Theorem 5 (main text). *There exists an activation operation $\mathcal{E}_{C \to AB} \in \mathcal{O}$ creating an SSR-entangled state σ_{AB} from ρ_C if and only if $\rho_C \notin \mathcal{S}$.*

Moreover, $\mathcal{E}$ can be taken to be any of the unitary operations described in Lemma 6, as long as all of the parameters r_i, t_i are non-vanishing.

Proof. We first prove that any particle-separable initial state results in no SSR-entanglement. This follows from a more general observation: any bipartite particle-separable state ρ_{AB} also SSR-separable. (This was stated in the two-particle case in Ref. [243].) As in the proof of Theorem D.4, a particle-separable bipartite state $|\psi\rangle_{AB}$ can be regarded as an effective two-mode state – taking a and b as linear combinations of the modes in A and

B respectively, we have

$$
\begin{aligned}
|\psi\rangle_{AB} &= \left(a^\dagger + b^\dagger\right)^N |0\rangle_A |0\rangle_B \\
&= \sum_{N_A} r_{N_A} |N_A\rangle_A |N - N_A\rangle_B ,
\end{aligned}
\tag{D.48}
$$

where the r_{N_A} are unimportant coefficients. It is immediate from this expression that $P_{N_A} \otimes P_{N-N_A} |\psi\rangle_{AB}$ is separable for all N_A. Since every particle-separable state is a convex combination of pure particle-separable states, the result follows for all mixed free states. So if ρ_C is a particle-separable state, then for any $\mathcal{E}_{C\to AB} \in \mathcal{O}$, $\mathcal{E}_{C\to AB}(\rho_C)$ is also particle-separable, and hence SSR-separable in the A/B partition.

Conversely, we prove that any unitary operation as in Lemma 6 with $r_i, t_i \neq 0 \ \forall i$ is sufficient to activate SSR-entanglement from PE. The simplest case – with a pure state and a "non-polarising beam-splitter", $r_i = r \ \forall i$ – was proven in Ref. [138]. Let us first argue that this extends to mixed states.

Suppose that the output state σ_{AB} is SSR-separable, so that each $(P_{N_A} \otimes P_{N_B})\sigma_{AB}(P_{N_A} \otimes P_{N_B})$ is separable. As shown in Ref. [138], the entanglement structure of $(P_{N_A} \otimes P_{N_B})\sigma_{AB}(P_{N_A} \otimes P_{N_B})$ is equivalent to $\rho_{N_A:N_B}^{\bullet(N_A+N_B)}$, in which the first-quantised form of the input state is partitioned into N_A versus N_B particles. Hence $\rho^{\bullet(N)}$ (with $N = N_A + N_B$) is bi-separable with respect to this partition, i.e.,

$$
\rho^{\bullet(N)} = \sum_i \lambda_i \, |\phi_i\rangle\langle\phi_i|_{N_A} \otimes |\chi_i\rangle\langle\chi_i|_{N_B} ,
\tag{D.49}
$$

where $|\phi_i\rangle \in \mathcal{H}_1^{\otimes N_A}$, $|\chi_i\rangle \in \mathcal{H}_1^{\otimes N_B}$, $\lambda_i \geq 0$. Since $\rho^{\bullet(N)}$ has support in the symmetric subspace $\mathcal{H}_N$, we must have $|\phi_i\rangle_{N_A} |\chi_i\rangle_{N_B} \in \mathcal{H}_N \ \forall i$. But any bi-separable symmetric pure state must also be fully separable. Therefore $|\phi_i\rangle_{N_A} |\chi_i\rangle_{N_B} = |\psi_i\rangle^{\otimes N}$, so $\rho^{\bullet(N)}$ is particle-separable.

Finally, we extend to the case of general r_i. Via a straightforward generalisation of the argument from Ref. [138], we find the output of the activation taking a Fock state $|\boldsymbol{n}\rangle$ as input – the details are in Appendix D.11. Denote by $|\xi\rangle_{AB}$ the output of activating $|\boldsymbol{n}\rangle$ with beam-splitter parameters $r_i = 1/\sqrt{2} \ \forall i$, and similarly denote by $|\eta\rangle_{AB}$ the output obtained with some arbitrary set of r_i. From (D.116) with two parties and $\alpha_{Ai} = r_i$, $\alpha_{Bi} =$

t_i, we have

$$
(P_{N_A} \otimes P_{N_B})|\eta\rangle_{AB}
$$

$$
= \binom{N}{N_A}^{1/2} \binom{N}{n}^{-1/2} \sum_{\substack{n_A \\ \sum_i n_{Ai} = N_A \\ n_{Bi} = n_i - n_{Ai}}} \binom{N_A}{n_A}^{1/2} \binom{N_B}{n_B}^{1/2}
$$

$$
\left[\prod_i r_i^{n_{Ai}} t_i^{n_{Bi}} \right] |n_A\rangle_A |n_B\rangle_B. \tag{D.50}
$$

It is clear from this expression that $|\eta\rangle$ can be obtained from $|\xi\rangle$ by application of the local operators $L_A \otimes L_B$, where

$$
L_A = \sum_{n_A} \left[\prod_i (\sqrt{2} r_i)^{n_{Ai}} \right] |n_A\rangle\langle n_A|,
$$

$$
L_B = \sum_{n_B} \left[\prod_i (\sqrt{2} t_i)^{n_{Bi}} \right] |n_B\rangle\langle n_B|. \tag{D.51}
$$

Since these operators are independent of the choice of initial Fock state, the same relationship holds for any input state – that is, the output from an arbitrary set of beam-splitters can be obtained by applying $L_A \otimes L_B$ to the output from a set of balanced beam-splitters. As long as $r_i, t_i \neq 0 \, \forall i$, these operators are invertible. The application of invertible local operators to a bipartite state does not change its Schmidt number [214]. This proves that the faithfulness of activation from a set of arbitrary non-trivial beam-splitters is equivalent to activation from balanced beam-splitters. $\square$

Theorem 6 (main text). *For any activation $\mathcal{E}_{C \to AB} \in \mathcal{O}$, $E_{SSR}^D(\mathcal{E}_{C \to AB}[\rho_C]) \leq M_{PE}^D(\rho_C)$.*

Proof. Let τ be the closest particle-separable state to ρ according to the measure D, then

$$
M_{\text{PE}}^D(\rho) = D(\rho, \tau) \tag{D.52}
$$

$$
\geq D(\mathcal{E}_{C \to AB}(\rho_C), \mathcal{E}_{C \to AB}(\tau_C)) \tag{D.53}
$$

$$
= D(\sigma_{AB}, \mathcal{E}_{C \to AB}(\tau_C)) \tag{D.54}
$$

$$
\geq D(\Phi_A \otimes \Phi_B(\sigma_{AB}), \Phi_A \otimes \Phi_B \circ \mathcal{E}_{C \to AB}(\tau_C)) \tag{D.55}
$$

$$
\geq E_{SSR}^D(\sigma_{AB}). \tag{D.56}
$$

The first two inequalities use the contractivity of D under channels. The final inequality uses the fact that τ is free, so that $\Phi_A \otimes \Phi_B \circ \mathcal{E}_{C \to AB}(\tau_C)$ is separable, but not in general the closest separable state to σ_{AB}. $\square$

Theorem 7 (main text). *For any (convex) entanglement measure E, the quantity de-*

fined as

$$M_{PE}^E(\rho) := \sup_{\mathcal{E}_{C\to AB}\in\mathcal{O}} E_{SSR}\left(\mathcal{E}_{C\to AB}[\rho_C]\right) \tag{D.57}$$

where the supremum is over all deterministic particle-separable operations, is a (convex) measure of PE.

Proof. The faithfulness of the measure is the content of Theorem 5. Deterministic monotonicity follows immediately from the definition and the fact that the set of operations $\mathcal{O}$ is closed under composition. Non-deterministic (strong) monotonicity states that $M_{\mathrm{PE}}^E(\rho)$ does not increase on average,

$$\sum_i p_i M_{\mathrm{PE}}^E(\sigma_i) \le M_{\mathrm{PE}}^E(\rho) \tag{D.58}$$

where $\Lambda_i(\rho) = p_i\sigma_i$ and $\{\Lambda_i\}_i \in \mathcal{O}$. From the definition (D.57), we have, for every activating channel $\mathcal{E}_{C\to AB} \in \mathcal{O}$,

$$M_{\mathrm{PE}}^E(\rho) \ge E_{SSR}\left(\mathcal{E}_{C\to AB}[\rho_C]\right). \tag{D.59}$$

We now continue to prove strong monotonicity by contradiction, showing that a violation of strong monotonicity (D.58), implies a violation of (D.59). If strong monotonicity (D.58) is violated, then there must exist a set of operations $\mathcal{E}_{i,C\to AB} \in \mathcal{O}$ such that the following is true:

$$M_{\mathrm{PE}}^E(\rho) < \sum_i p_i E_{SSR}\left(\mathcal{E}_{i,C\to AB}[\sigma_{i,C}]\right). \tag{D.60}$$

We now invoke a general property of entanglement measures (and SSR-entanglement measures), namely monotonicity under the partial trace over a subsystem. We split B into two subsystems B_1, B_2, in which B_2 contains a classical flag. Then, for any ensemble of state ρ_{i,AB_1} with probabilities p_i,

$$E_{SSR}\left(\sum_i p_i\rho_{i,AB_1}\otimes|i\rangle\langle i|_{B_2}\right) \ge \sum_i p_i E_{SSR}\left(\rho_{i,AB_1}\right). \tag{D.61}$$

Applying this to (D.60), we obtain

$$M_{\mathrm{PE}}^E(\rho) < E_{SSR}\left(\sum_i p_i\mathcal{E}_{i,C\to AB_1}[\sigma_{i,C}]\otimes|i\rangle\langle i|_{B_2}\right)$$

$$< E_{SSR}\left(\sum_i \mathcal{E}_{i,C\to AB_1}[\Lambda_{i,C}(\rho_C)]\otimes|i\rangle\langle i|_{B_2}\right). \tag{D.62}$$

Note that the operations appearing on the right-hand side above can be combined into a single operation $\mathcal{F}_{C \to AB_1 B_2} \in \mathcal{O}$, which is performed by first applying $\{\Lambda_i\}_i$, storing the outcome i in a classical flag, and then conditionally applying $\mathcal{E}_i$. Thus,

$$M_{\mathrm{PE}}^E(\rho) < E_{SSR}\left(\mathcal{F}_{C \to AB_1 B_2}[\rho_C]\right). \tag{D.63}$$

The above is a direct contradiction of (D.59), thus establishing that M_{PE} is a strong monotone for any entanglement monotone E_{SSR}.

We now continue by showing convexity:

$$M_{\mathrm{PE}}^E\left(\sum_i p_i \rho_i\right) \leq \sum_i p_i M_{\mathrm{PE}}(\rho_i). \tag{D.64}$$

From the definition of M_{PE}, we have

$$
\begin{aligned}
M_{\mathrm{PE}}^E\left(\sum_i p_i \rho_i\right) &\leq \sup_{\mathcal{E}_{C \to AB} \in \mathcal{O}} \sum_i p_i E_{SSR}\left(\mathcal{E}_{C \to AB}[\rho_{i,C}]\right) \\
&\leq \sum_i p_i \left\{\sup_{\mathcal{E}_{C \to AB} \in \mathcal{O}} E_{SSR}\left(\mathcal{E}_{C \to AB}[\rho_{i,C}]\right)\right\} \\
&= \sum_i p_i M_{PE}^E(\rho_i).
\end{aligned}
\tag{D.65}
$$

where we have made use of the fact that taking the supremum over each term in the sum individually cannot give less than a single supremum. $\qquad\square$

D.9   Lower bound on PE measure from an entanglement criterion

In order to witness the entanglement present in the system a criterion of separability from [99] is used, which is satisfied for all separable states,

$$1 \leq \frac{4 \mathrm{Var}\left(g_z \hat{S}_z^A + \hat{S}_z^B\right) \mathrm{Var}\left(g_y \hat{S}_y^A + \hat{S}_y^B\right)}{\left(|g_z g_y| \left|\left\langle \hat{S}_x^A \right\rangle\right| + \left|\left\langle \hat{S}_x^B \right\rangle\right|\right)^2}, \tag{D.66}$$

where $\mathrm{Var}(\cdot)$ denotes the variance and $g_{(y,z)}$ are real parameters that can be optimised over. The z-component of the spin in regions A, B is defined as $\hat{S}_z^{(A,B)} :=$

$\frac{1}{2\eta_{\text{eff}}^{(A,B)}}\left(\hat{N}_1^{(A,B)} - \hat{N}_2^{(A,B)}\right)$ where $1, 2$ correspond to the internal degree of freedom of the atom and $\eta_{\text{eff}}^{(A,B)}$ accounts for finite spatial resolution in the detection of the BEC. Other spin components, e.g. $\hat{S}_x^{(A,B)}$ and $\hat{S}_y^{(A,B)}$, can be measured by applying appropriate spin rotations before detection. In the following we will show that this condition of separability (D.66), can be rewritten as an entanglement witness.

Taking the root of equation (D.66) and collecting the terms,

$$
0 \le 2\sqrt{\text{Var}\left(g_z\hat{S}_z^A + \hat{S}_z^B\right)\text{Var}\left(g_y\hat{S}_y^A + \hat{S}_y^B\right)}
$$
$$
- \left(|g_zg_y|\left|\left\langle\hat{S}_x^A\right\rangle\right| + \left|\left\langle\hat{S}_x^B\right\rangle\right|\right)
$$
$$
0 \le \text{Var}\left(g_z\hat{S}_z^A + \hat{S}_z^B\right) + \text{Var}\left(g_y\hat{S}_y^A + \hat{S}_y^B\right)
$$
$$
- \left(|g_zg_y|\left\langle\hat{S}_x^A\right\rangle + \left\langle\hat{S}_x^B\right\rangle\right)
$$
$$
0 \le \text{Var}\left(g_z\hat{S}_z^A + \hat{S}_z^B\right) + \text{Var}\left(g_y\hat{S}_y^A + \hat{S}_y^B\right)
$$
$$
- \left\langle|g_zg_y|\,\hat{S}_x^A + \hat{S}_x^B\right\rangle, \tag{D.67}
$$

where in the second line we have applied the inequality between the geometric and arithmetic mean and removed some of the absolute signs in the third term. We can simplify notation by defining component spin operators $\hat{S}_z^+ := g_z\hat{S}_z^A + \hat{S}_z^B$, $\hat{S}_y^+ := g_y\hat{S}_y^A + \hat{S}_y^B$ and $\hat{S}_x^+ := |g_zg_y|\,\hat{S}_x^A + \hat{S}_x^B$,

$$
\text{Var}\left(\hat{S}_z^+\right) + \text{Var}\left(\hat{S}_y^+\right) - \left\langle\hat{S}_x^+\right\rangle \ge 0. \tag{D.68}
$$

We now relate this to an entanglement witness observable. For any state ρ, let

$$
W_\rho := \left(\hat{S}_z^+ - \left\langle\hat{S}_z^+\right\rangle_\rho\right)^2 + \left(\hat{S}_y^+ - \left\langle\hat{S}_y^+\right\rangle_\rho\right)^2 - \hat{S}_x^+. \tag{D.69}
$$

To check that this is a valid entanglement witness, let σ be any separable state. Using $\langle(X - x_0)^2\rangle = V(X) + (x_0 - \langle X\rangle)^2$, from (D.68) we have

$$
\text{Tr}[\sigma W_\rho] = \left\langle\left(\hat{S}_z^+ - \left\langle\hat{S}_z^+\right\rangle_\rho\right)^2\right\rangle_\sigma
$$
$$
+ \left\langle\left(\hat{S}_y^+ - \left\langle\hat{S}_y^+\right\rangle_\rho\right)^2\right\rangle_\sigma - \left\langle\hat{S}_x^+\right\rangle_\sigma \ge 0. \tag{D.70}
$$

Note that when the ρ defining W_ρ is chosen to be the same as the state being measured, the expectation value $\text{Tr}[\rho W_\rho]$ equals the left-hand side of (D.68).

Now we have defined an entanglement witness, we can relate such a quantity to a commonly used measure of entanglement defined as the trace distance to the set of separable states,

$$M_{\mathrm{PE}}^{\mathrm{Tr}}(\rho) := \min_{\sigma \in \text{ sep.}} \max_{0 \leq P \leq \mathbb{1}} \mathrm{Tr}\left[P(\sigma - \rho)\right], \tag{D.71}$$

where P is hermitian. This is by no means the only entanglement measure that can be related to our witness [42] but provides a convenient form. As both P and σ vary within compact convex sets, and the trace distance is concave for fixed σ and convex for fixed P, we can make use of the minimax theorem [177] to obtain

$$M_{\mathrm{PE}}^{\mathrm{Tr}}(\rho) = \max_{0 \leq P \leq \mathbb{1}} \min_{\sigma \in \text{ sep.}} \mathrm{Tr}\left[P(\sigma - \rho)\right]. \tag{D.72}$$

Now in order to write this measure in terms of the entanglement witness W_ρ we choose a particular P:

$$P = W_\rho' + c\mathbb{1}, \tag{D.73}$$

where c is a constant and $W_\rho' = W_\rho/\mathcal{N}$ is a normalised witness with the factor $\mathcal{N}$ to be determined later. The constants must be chosen appropriately such that $0 \leq P \leq \mathbb{1}$. This condition is equivalent to

$$-c\mathbb{1} \leq W_\rho' \leq (1-c)\mathbb{1}, \tag{D.74}$$

which implies that $0 < c < 1$ since the witness can take values of both signs. Then we have

$$\begin{aligned}
M_{\mathrm{PE}}^{\mathrm{Tr}}(\rho) &\geq \min_{\sigma \in \text{ sep.}} \left[\mathrm{Tr}\left[W_\rho'(\sigma - \rho)\right] + c\,\mathrm{Tr}\left[\mathbb{1}(\sigma - \rho)\right] \right] \\
&\geq -\mathrm{Tr}\left[W_\rho'\rho\right] + \min_{\sigma \in \text{ sep.}} \mathrm{Tr}\left[W_\rho'\sigma\right] \\
&\geq -\mathrm{Tr}\left[W_\rho'\rho\right],
\end{aligned} \tag{D.75}$$

where we have used the fact that $\min_{\sigma \in \text{ sep.}} \mathrm{Tr}\left[W_\rho'\sigma\right] \geq 0$.

We optimise c and $\mathcal{N}$ to obtain the maximal lower bound on $M_{\mathrm{PE}}^{\mathrm{Tr}}(\rho)$ subject to normalisation constraints. We start by writing down the range of values taken by the witness,

$$W_\rho^- \leq \langle W_\rho \rangle \leq W_\rho^+, \tag{D.76}$$

where W_ρ^- and W_ρ^+ are the minimum and maximum eigenvalues of W_ρ. The objective is to make $W_\rho^-/\mathcal{N}$ as negative as possible. Using equation (D.74), for given c we want the

minimum value of $\mathcal{N}$ such that $\mathcal{N} \geq -W^-/c$ and $\mathcal{N} \geq W^+/(1-c)$ are both true. We therefore want to choose the normalisation $\mathcal{N}(c)$ such that

$$\mathcal{N}(c) = \max \left\{ \frac{-W^-}{c}, \frac{W^+}{1-c} \right\}. \tag{D.77}$$

We can see that the minimum value of $\mathcal{N}(c)$ occurs (for a certain constant c^*), when these two terms are equal. We have

$$c^* = \frac{W_\rho^-}{W_\rho^- - W_\rho^+} \tag{D.78}$$

and substituting this back into equation (D.77) gives us the normalisation constant,

$$\mathcal{N}(c^*) = W_\rho^+ - W_\rho^-. \tag{D.79}$$

So the bound on the entanglement measure can therefore be written as,

$$M_{\text{PE}}^{\text{Tr}}(\rho) \geq \frac{-1}{W_\rho^+ - W_\rho^-} \, \text{Tr}\left[W_\rho \rho\right]. \tag{D.80}$$

We continue by calculating upper and lower bounds for W_ρ^+ and W_ρ^- respectively. Starting with W_ρ^- we lower bound the product of the variances in the first line of equation (D.67) using the Robertson uncertainty relation,

$$
\begin{aligned}
\text{Var}\left(g_z \hat{S}_z^A + \hat{S}_z^B\right) &\text{Var}\left(g_y \hat{S}_y^A + \hat{S}_y^B\right) \\
&\geq \frac{1}{4}\left|\left\langle g_z g_y \left[\hat{S}_z^A, \hat{S}_y^A\right] + \left[\hat{S}_z^B, \hat{S}_y^B\right]\right\rangle\right|^2 \\
&= \frac{1}{4}\left|\left\langle -i g_z g_y \hat{S}_x^A - i \hat{S}_x^B \right\rangle\right|^2 \\
&= \frac{1}{4}\left\langle g_z g_y \hat{S}_x^A + S_x^B \right\rangle^2,
\end{aligned} \tag{D.81}
$$

where we have used the standard spin commutator relations. This can now be substituted back into the first line of equation (D.67) to lower bound W_ρ^- where again we write the second term as a single expectation value,

$$
\begin{aligned}
W_\rho^- &\geq \min_\sigma \left[\left|\left\langle g_z g_y \hat{S}_x^A + \hat{S}_x^B \right\rangle_\sigma\right| - \left\langle |g_z g_y| \, \hat{S}_x^A + \hat{S}_x^B \right\rangle_\sigma\right] \\
&\geq 0 - \max_\sigma \left\langle |g_z g_y| \, \hat{S}_x^A + \hat{S}_x^B \right\rangle_\sigma.
\end{aligned} \tag{D.82}
$$

The spin operators take their maximal value when all the particles are in internal mode

1, $\max \hat{S}^{(A,B)} = \frac{1}{2\eta_{\text{eff}}^{(A,B)}} N^{(A,B)}$.

$$W_\rho^- \geq -\frac{1}{2} \left(\frac{|g_z g_y| N^A}{\eta_{\text{eff}}^A} + \frac{N^B}{\eta_{\text{eff}}^B} \right), \tag{D.83}$$

providing us with a lower bound on W_ρ^-. We now move onto upper bounding W_ρ^+. We can start by upper bounding the variance terms in the last line of equation (D.67). This can be achieved by utilizing Popoviciu's inequality [194],

$$\begin{aligned} \text{Var}\left(g_z \hat{S}_z^A + \hat{S}_z^B \right) &\leq \frac{1}{4} \left(\lambda_{\max}\left(g_z \hat{S}_z^A + \hat{S}_z^B \right) - \lambda_{\min}\left[g_z \hat{S}_z^A + \hat{S}_z^B \right] \right)^2 \\ &= \lambda_{\max}\left[g_z \hat{S}_z^A + \hat{S}_z^B \right]^2 \\ &= \left(|g_z| \lambda_{\max}\left[\hat{S}_z^A \right] + \lambda_{\max}\left[\hat{S}_z^B \right] \right)^2 \\ &= \frac{1}{4} \left(\frac{|g_z| N^A}{\eta_{\text{eff}}^A} + \frac{N^B}{\eta_{\text{eff}}^B} \right)^2 \end{aligned} \tag{D.84}$$

where $\lambda_{\max}[A], \lambda_{\min}[A]$ are the maximum and minimum eigenvalues of the operator A, respectively, and in last line we have again used the fact that the value is maximised when all the particles are in the same internal mode. Substituting the above into the last line of equation (D.67) and maximising over each term individually results in,

$$\begin{aligned} W_\rho^+ &\leq \frac{1}{4} \left(\frac{|g_z| N^A}{\eta_{\text{eff}}^A} + \frac{N^B}{\eta_{\text{eff}}^B} \right)^2 + \frac{1}{4} \left(\frac{|g_y| N^A}{\eta_{\text{eff}}^A} + \frac{N^B}{\eta_{\text{eff}}^B} \right)^2 \\ &\quad - \min_\sigma \left\langle |g_z g_y| \hat{S}_x^A + \hat{S}_x^B \right\rangle_\sigma \\ &\leq \frac{1}{4} \left(\frac{|g_z| N^A}{\eta_{\text{eff}}^A} + \frac{N^B}{\eta_{\text{eff}}^B} \right)^2 + \frac{1}{4} \left(\frac{|g_y| N^A}{\eta_{\text{eff}}^A} + \frac{N^B}{\eta_{\text{eff}}^B} \right)^2 \\ &\quad + \frac{1}{2} \left(\frac{|g_z g_y| N^A}{\eta_{\text{eff}}^A} + \frac{N^B}{\eta_{\text{eff}}^B} \right). \end{aligned} \tag{D.85}$$

Now we have bounded both the maximum and minimum values the witness can take, we can bound the normalisation $\mathcal{N}$ from equation (D.79) and therefore bound the entanglement measure with a normalised witness,

$$\begin{aligned} M_{\text{PE}}^{\text{Tr}}(\rho) \geq -&\left[\frac{1}{4} \left(\frac{|g_z| N^A}{\eta_{\text{eff}}^A} + \frac{N^B}{\eta_{\text{eff}}^B} \right)^2 + \frac{1}{4} \left(\frac{|g_y| N^A}{\eta_{\text{eff}}^A} + \frac{N^B}{\eta_{\text{eff}}^B} \right)^2 \right. \\ &\left. + \left(\frac{|g_z g_y| N^A}{\eta_{\text{eff}}^A} + \frac{N^B}{\eta_{\text{eff}}^B} \right) \right]^{-1} \text{Tr}\left[W_\rho \rho \right]. \end{aligned} \tag{D.86}$$

D.10 Non-classicality

Theorem. *Every number-diagonal (ND) classical state is particle-separable.*

Proof. If ρ is classical and ND, then

$$\rho = \int \mathrm{d}^{2n}\boldsymbol{\alpha}\, P(\boldsymbol{\alpha})\Phi(|\boldsymbol{\alpha}\rangle\langle\boldsymbol{\alpha}|), \tag{D.87}$$

with $P(\boldsymbol{\alpha}) \geq 0$. Hence it is sufficient to prove the claim for all $\Phi(|\boldsymbol{\alpha}\rangle\langle\boldsymbol{\alpha}|)$. For any multi-mode coherent state $|\boldsymbol{\alpha}\rangle$, there exists a passive linear unitary U that brings all the particles into a single mode: $U|\boldsymbol{\alpha}\rangle = |\bar{\alpha}\rangle|0\rangle^{\otimes(n-1)}$, where $|\bar{\alpha}|^2 = \sum_{i=1}^n |\alpha_i|^2$. Since this unitary is number-conserving, it commutes with Φ, so

$$U\Phi(|\boldsymbol{\alpha}\rangle\langle\boldsymbol{\alpha}|)U^\dagger = \Phi\left(U|\boldsymbol{\alpha}\rangle\langle\boldsymbol{\alpha}|U^\dagger\right) \tag{D.88}$$

$$= \Phi\left(|\bar{\alpha}\rangle\langle\bar{\alpha}| \otimes |0\rangle\langle 0|^{\otimes(n-1)}\right) \tag{D.89}$$

$$= \Phi(|\bar{\alpha}\rangle\langle\bar{\alpha}|) \otimes |0\rangle\langle 0|^{\otimes(n-1)} \tag{D.90}$$

$$= \sum_{k=0}^\infty \frac{e^{-|\bar{\alpha}|^2}|\bar{\alpha}|^{2k}}{k!}|k\rangle\langle k| \otimes |0\rangle\langle 0|^{\otimes(n-1)}, \tag{D.91}$$

which is particle-separable. $\qquad\square$

Theorem 8 (main text). *Two copies $\rho^{\otimes 2}$ of a number-bounded state ρ are particle-separable if and only if ρ is the vacuum.*

Proof. Let both ρ and $\rho^{\otimes 2}$ be free with bounded particle number, and we decompose $\rho = \sum_{N=0}^{N_0} p_N \rho^{(N)}$. Then

$$\rho^{\otimes 2} = \sum_{N,N'=0}^{N_0} p_N p_{N'} \rho^{(N)} \otimes \rho^{(N')}. \tag{D.92}$$

The maximal number component of this state is $p_{N_0}^2 \rho^{(N_0)} \otimes \rho^{(N_0)}$, where $p_{N_0} \neq 0$ by assumption. This component must be particle-separable, thus must be obtainable by mixtures of the form $\sum_i p_i U_i |2N_0, 0, 0, \ldots\rangle\langle 2N_0, 0, 0, \ldots|U_i^\dagger$, where the U_i are passive linear. Now this state has exactly N_0 particles on each of the two parties, and so the same must be true for every term in the sum. In other words, for each i, $U_i|2N_0, 0\rangle = (V_i|N_0\rangle)(W_i|N_0\rangle)$ with pair of additional passive linear unitaries V_i, W_i acting on each subsystem. It is easily seen that this is impossible unless $N_0 = 0$. $\qquad\square$

Theorem 9 (main text). *Two copies $\Phi(|\psi\rangle\langle\psi|)^{\otimes 2}$ of a pseudo-pure state are particle-separable if and only if $|\psi\rangle$ is classical.*

Proof. We first show that the activation of an arbitrary pure state $|\psi\rangle$ into SSR-entanglement is exactly the same as for the pseudo-pure state $\Phi(|\psi\rangle\langle\psi|)$. Let Φ_{AB} be the joint dephasing operator with respect to the *total* number over two parties A, B. This operation is already implemented by dephasing with respect to local number, so that $(\Phi_A \otimes \Phi_B) = (\Phi_A \otimes \Phi_B) \circ \Phi_{AB}$. We use this to connect the SSR-entanglement activated by a unitary $\mathcal{U} \in \mathcal{O}$ from $|\psi\rangle\langle\psi|$ to that activated from $\Phi(|\psi\rangle\langle\psi|)$:

$$(\Phi_A \otimes \Phi_B) \circ \mathcal{U}\big(|\psi\rangle\langle\psi|_A \otimes |0\rangle\langle0|_B\big)$$

$$= (\Phi_A \otimes \Phi_B) \circ \Phi_{AB} \circ \mathcal{U}\big(|\psi\rangle\langle\psi|_A \otimes |0\rangle\langle0|_B\big) \tag{D.93}$$

$$= (\Phi_A \otimes \Phi_B) \circ \mathcal{U} \circ \Phi_{AB}\big(|\psi\rangle\langle\psi|_A \otimes |0\rangle\langle0|_B\big) \tag{D.94}$$

$$= (\Phi_A \otimes \Phi_B) \circ \mathcal{U}\big(\Phi_A[|\psi\rangle\langle\psi|_A] \otimes |0\rangle\langle0|_B\big), \tag{D.95}$$

where we have used the fact that $\mathcal{U}$ is number-conserving, so $[\mathcal{U}, \Phi_{AB}] = 0$, and the last line holds because B contains no particles.

Now let $|\psi\rangle$ be activated by $\mathcal{U}$ consisting of a set of non-trivial beam-splitters into $|\phi\rangle_{AB}$. Then we can write $|\phi\rangle_{AB} = \sum_{k,l} |\phi_{k,l}\rangle_{AB} := \sum_{k,l} P_{k,A} P_{l,B} |\phi\rangle_{AB}$. If two copies of $|\psi\rangle$ are activated in the same way in parallel, then the output state is $|\phi\rangle^{\otimes 2} = |\phi\rangle_{A_1 B_1} |\phi\rangle_{A_2 B_2}$. Given that $\Phi(|\psi\rangle\langle\psi|)^{\otimes 2}$ is particle-separable, Theorem 5 says that the projection of the activated state onto local particle number must be unentangled – so there exist (unnormalised) $|a^{n,m}\rangle_{A_1 A_2}$, $|b^{n,m}\rangle_{B_1 B_2}$ such that, for each n, m,

$$P_{n,A} P_{m,B} |\phi\rangle_{A_1 B_1} |\phi\rangle_{A_2 B_2} = |a^{n,m}\rangle_{A_1 A_2} |b^{n,m}\rangle_{B_1 B_2}. \tag{D.96}$$

Applying the projector $P_{k,A_1} P_{l,B_1}$ onto local numbers in the first copy, we find

$$|\phi_{k,l}\rangle_{A_1 B_1} |\phi_{n-k,m-l}\rangle_{A_2 B_2}$$

$$= \big(P_{k,A_1} |a^{n,m}\rangle_{A_1 A_2}\big) \big(P_{l,B_1} |b^{n,m}\rangle_{B_1 B_2}\big). \tag{D.97}$$

Both sides of the above equation must be separable with respect to both the $A_1 A_2 / B_1 B_2$ and $A_1 B_1 / A_2 B_2$ partitions. Therefore there must exist (unnormalised) states $|a_k^{n,m}\rangle_{A_1}$, $|b_l^{n,m}\rangle_{B_1}$ such that

$$|\phi_{k,l}\rangle_{A_1 B_1} = |a_k^{n,m}\rangle_{A_1} |b_l^{n,m}\rangle_{B_1}. \tag{D.98}$$

The left-hand side of the above is independent of n and m, so the same must be true of the states on the right – removing these labels, we obtain

$$|\phi_{k,l}\rangle_{A_1 B_1} = |a_k\rangle_{A_1} |b_l\rangle_{B_1}. \tag{D.99}$$

Summing over k and l, we see that $|\phi_{k,l}\rangle_{A_1 B_1} = (\sum_k |a_k\rangle_{A_1})(\sum_l |b_l\rangle_{B_1})$ is separable. From

the result in quantum optics saying that all non-classical states are activated into entangled states, it follows that $|\psi\rangle$ must be classical. $\qquad\qquad\square$

In the following, the vacuum state of any number of modes will be denoted $|0\rangle$. The primitive system S under consideration has d modes, and we denote k copies of S by S^k.

The proof of Theorem 10 relies on the following result, which is of the "de Finetti" type [53].

Theorem. *Let $\rho_{[m]}$ be an exchangeable (i.e., permutation-symmetric) state of N particles on m modes that is also particle-separable. Denote by $\rho_{[l]}$ the reduced state of any subset of $l \leq m$ modes. Then there exists a classical l-mode state $\sigma_{[l]}$ such that*

$$D_{\mathrm{Tr}}(\rho_{[l]}, \sigma_{[l]}) \leq \frac{l}{m}, \qquad\qquad (D.100)$$

Proof. Since $\rho_{[m]}$ is particle separable, there is a probability distribution q_λ and a set of single-particle creation operators $c_\lambda^\dagger$ such that

$$\rho_{[m]} = \sum_\lambda \frac{q_\lambda}{N!} (c_\lambda^\dagger)^N |0\rangle\langle 0| c_\lambda^N. \qquad\qquad (D.101)$$

We decompose $c_\lambda^\dagger = \alpha_\lambda a_\lambda^\dagger + \alpha_\lambda' a_\lambda'^\dagger$, where $|\alpha_\lambda|^2 + |\alpha_\lambda'|^2 = 1$, a_λ acts on modes $1, \ldots, l$, and a_λ' acts on modes $l+1, \ldots, m$. Using the binomial expansion for $(c_\lambda^\dagger)^N$ and tracing out modes $l+1, \ldots, m$, we have

$$\begin{aligned}
\rho_{[l]} &= \mathrm{Tr}_{l+1,\ldots,m}\, \rho_{[m]} \\
&= \sum_{n=0}^N \frac{1}{N!} \binom{N}{n}^2 |\alpha_\lambda|^{2n} |\alpha_\lambda'|^{2(N-n)} (a_\lambda^\dagger)^n |0\rangle\langle 0| a_\lambda^n \\
&= \sum_{n=0}^N \binom{N}{n} |\alpha_\lambda|^{2n} (1 - |\alpha_\lambda|^2)^{N-n} |n^{(\lambda)}\rangle\langle n^{(\lambda)}| \\
&= \sum_{n=0}^N b_\lambda(n) |n^{(\lambda)}\rangle\langle n^{(\lambda)}|, \qquad\qquad (D.102)
\end{aligned}$$

where b_λ is the binomial distribution with N trials and $p = |\alpha_\lambda|^2$, and $|n^{(\lambda)}\rangle := \frac{1}{\sqrt{n!}} (a_\lambda^\dagger)^n |0\rangle$.

Now we use a result on the Poisson distribution as a limit case of the binomial distribution. For a binomial $b(n)$ and Poisson $\pi(n)$ with the same mean μ, it is well known that $b \to \pi$ in the limit of large N. In fact, a stronger result [111](Eq. 4) says that

$$D_{\mathrm{Tr}}(b, \pi) \leq p = \frac{\mu}{N}, \qquad\qquad (D.103)$$

where D_{Tr} here is the classical version of the trace distance.

Let π_λ be the Poisson distribution with mean $\mu_k = N|\alpha_\lambda|^2$, and define

$$\sigma_{[l]} := \sum_\lambda q_\lambda \sum_{n=0}^\infty \pi_\lambda(n) \left|n^{(\lambda)}\right\rangle\!\left\langle n^{(\lambda)}\right|. \tag{D.104}$$

Note that $\sigma_{[l]}$ is classical since it can be written in the form

$$\sigma_{[l]} = \sum_\lambda q_\lambda \Phi(|\psi_\lambda\rangle\langle\psi_\lambda|), \tag{D.105}$$

$$|\psi_\lambda\rangle := \sum_{n=0}^\infty \sqrt{\pi_\lambda(n)} \left|n^{(\lambda)}\right\rangle, \tag{D.106}$$

where $|\psi_\lambda\rangle$ is a coherent state with mean particle number μ_λ. It follows that

$$
\begin{aligned}
D_{\mathrm{Tr}}(\rho_{[l]}, \sigma_{[l]}) &= \frac{1}{2}\left\|\sum_{\lambda,n} q_\lambda[b_\lambda(n) - \pi_\lambda(n)]\left|n^{(\lambda)}\right\rangle\!\left\langle n^{(\lambda)}\right|\right\|_1 \\
&\le \frac{1}{2}\sum_{\lambda,n} q_\lambda \left\|[b_\lambda(n) - \pi_\lambda(n)]\left|n^{(\lambda)}\right\rangle\!\left\langle n^{(\lambda)}\right|\right\|_1 \\
&= \sum_\lambda q_\lambda \sum_n \frac{1}{2}|b_\lambda(n) - \pi_\lambda(n)| \\
&= \sum_\lambda q_\lambda D_{\mathrm{Tr}}(b_\lambda, \pi_\lambda) \\
&\le \sum_\lambda q_\lambda \frac{\mu_\lambda}{N},
\end{aligned} \tag{D.107}
$$

having used the triangle inequality and finally (D.103). Now $\sum_\lambda q_\lambda \mu_\lambda$ is the mean particle number in $\rho_{[l]}$, which by exchangeability is Nl/m. Therefore

$$D_{\mathrm{Tr}}(\rho_{[l]}, \sigma_{[l]}) \le \frac{l}{m}. \tag{D.108}$$

$\square$

Theorem 10 (main text). *Let ρ have finite mean particle number, $\mathrm{Tr}\!\left[\rho\hat{N}\right] < \infty$, and suppose that $\rho^{\otimes k}$ is particle-separable for some k. Then the trace-distance non-classicality of ρ is bounded by*

$$M_{\mathrm{NC}}^{\mathrm{Tr}}(\rho) \le \frac{1}{k}. \tag{D.109}$$

Consequently, $\rho^{\otimes k}$ is particle-separable for all k if and only if ρ is classical.

Proof. Let ρ contain d modes, so that $\rho^{\otimes k}$ contains $m = kd$ modes. Projecting onto the subspace of total particle number N results in the (normalised) state $P_{N,S^k}\rho^{\otimes k}P_{N,S^k}/p_N$,

which fulfils the assumptions of Theorem D.10. Therefore there exists a classical state σ_N of d modes such that

$$D_{\text{Tr}}\left(\frac{\text{Tr}_{S_2,\dots,S_k} P_{N,S^k}\rho^{\otimes k}P_{N,S^k}}{p_N}, \sigma_N\right) \leq \frac{d}{kd} = \frac{1}{k}. \tag{D.110}$$

Defining the classical state $\sigma := \sum_N p_N\sigma_N$, we have

$$
\begin{aligned}
D_{\text{Tr}}(\rho,\sigma) &= D_{\text{Tr}}\left(\sum_N \text{Tr}_{S_2,\dots,S_k} P_{N,S^k}\rho^{\otimes k}P_{N,S^k}, \sum_N p_N\sigma_N\right) \\
&\leq \sum_N p_N D_{\text{Tr}}\left(\frac{\text{Tr}_{S_2,\dots,S_k} P_{N,S^k}\rho^{\otimes k}P_{N,S^k}}{p_N}, \sigma_N\right) \\
&\leq \sum_N p_N \frac{1}{k} \\
&= \frac{1}{k},
\end{aligned}
\tag{D.111}
$$

having used convexity of D_{Tr}.

The final statement is an immediate application of this bound in the limit $k \to \infty$, using the fact that the set of classical states is closed in the trace-norm topology [17]. Conversely, it is enough to note that the set of classical states is closed under tensor products. $\qquad\square$

D.11 Unitary activation of Fock states

Here we generalise the main result of Ref. [138] to multiple modes and to general beam-splitters. We also present the results without much additional effort for arbitrary numbers of parties, although the rest of our work uses only the bipartite case. Let us first find the first-quantised form of an m-mode Fock state $|\boldsymbol{n}\rangle$, partitioned into sets of $N_A, N_B, \dots, N_Z$ particles, where $\sum_{K=A,B,\dots,Z} N_K = N := \sum_i n_i$. We have

$$|\boldsymbol{n}\rangle^{\bullet} = \binom{N}{\boldsymbol{n}}^{-1/2} \sum_{\Pi} \Pi \bigotimes_{i=0}^{m-1} |i\rangle^{\otimes n_i}, \tag{D.112}$$

where $\binom{N}{\boldsymbol{n}}$ is a multinomial coefficient and the sum runs over *distinct* permutations Π of $\bigotimes_{i=0}^{m-1} |i\rangle^{\otimes n_i}$. Dividing initially into N_A versus $N_{\bar{A}} = N - N_A$ particles, it may be verified

that

$$|\boldsymbol{n}\rangle^{\bullet} = \binom{N}{\boldsymbol{n}}^{-1/2} \sum_{\substack{\{n_{Ai}\}_i \\ \sum_i n_{Ai}=N_A}} \binom{N_A}{\boldsymbol{n_A}}^{1/2} \binom{N_{\bar{A}}}{\boldsymbol{n_{\bar{A}}}}^{1/2} |\boldsymbol{n_A}\rangle^{\bullet}_{N_A} |\boldsymbol{n_{\bar{A}}}\rangle^{\bullet}_{N_{\bar{A}}}, \tag{D.113}$$

where $n_{\bar{A}i} = n_i - n_{Ai}$. Recursively continuing the subdivision of $\bar{A}$ in this way, we obtain

$$|\boldsymbol{n}\rangle^{\bullet} = \binom{N}{\boldsymbol{n}}^{-1/2} \sum_{\substack{\{\boldsymbol{n_K}\}_K \\ \sum_i n_{Ki}=N_K \forall K \\ \sum_K n_{Ki}=n_i \forall i}} \bigotimes_K \binom{N_K}{\boldsymbol{n_K}}^{1/2} |\boldsymbol{n_K}\rangle^{\bullet}_{N_K}. \tag{D.114}$$

Next, we show how a Fock state on A is activated into a multipartite SSR-entangled state by mixing with vacuum modes on $B, \ldots, Z$ at a generalised beam splitter. Specifically, we take the beam-splitter U to have the action $a^{\dagger}_{Ai} \rightarrow \sum_K \alpha_{Ki} a^{\dagger}_{Ki}$ – a generalisation of Ref. [138], in which α_{Ki} was independent of i. Then

$$
\begin{aligned}
|\phi\rangle_{A\ldots Z} &:= U|\boldsymbol{n}\rangle_A |00\ldots\rangle_{B\ldots Z} \\
&= \prod_i \frac{1}{\sqrt{n_i!}} \left(\sum_K \alpha_{Ki} a^{\dagger}_{Ki} \right)^{n_i} |00\ldots\rangle_{A\ldots Z} \\
&= \prod_i \frac{1}{\sqrt{n_i!}} \sum_{\substack{\{n_{Ki}\} \\ \sum_K n_{Ki}=n_i \forall i}} \binom{n_i}{n_{Ai}, \ldots, n_{Zi}} \\
&\qquad \prod_K (\alpha_{Ki} a^{\dagger}_{Ki})^{n_{Ki}} |00\ldots\rangle_{A\ldots Z} \\
&= \sum_{\substack{\{\boldsymbol{n_K}\}_K \\ \sum_K n_{Ki}=n_i \forall i}} \left[\prod_i \binom{n_i}{n_{Ai}, \ldots, n_{Zi}}^{1/2} \right] \\
&\qquad \bigotimes_K \left[\prod_i \alpha^{n_{Ki}}_{Ki} \right] |\boldsymbol{n_K}\rangle_K. \tag{D.115}
\end{aligned}
$$

Conditioning on local particle number,

$$
\begin{aligned}
(P_{N_A} &\otimes \ldots \otimes P_{N_Z})|\phi\rangle_{A\ldots Z} \\
&= \sum_{\substack{\{\boldsymbol{n_K}\}_K \\ \sum_i n_{Ki}=N_K \forall K \\ \sum_K n_{Ki}=n_i \forall i}} \left[\prod_i \binom{n_i}{n_{Ai},\ldots,n_{Zi}}^{1/2} \right] \bigotimes_K \left[\prod_i \alpha_{Ki}^{n_{Ki}} \right] |\boldsymbol{n_K}\rangle_K \\
&= \left[\frac{\prod_i n_i!}{\prod_K N_K!} \right]^{1/2} \sum_{\substack{\{\boldsymbol{n_K}\}_K \\ \sum_i n_{Ki}=N_K \forall K \\ \sum_K n_{Ki}=n_i \forall i}} \bigotimes_K \binom{N_K}{\boldsymbol{n_K}}^{1/2} \left[\prod_i \alpha_{Ki}^{n_{Ki}} \right] |\boldsymbol{n_K}\rangle_K \\
&= \binom{N}{N_A,\ldots,N_Z}^{1/2} \binom{N}{\boldsymbol{n}}^{-1/2} \sum_{\substack{\{\boldsymbol{n_K}\}_K \\ \sum_i n_{Ki}=N_K \forall K \\ \sum_K n_{Ki}=n_i \forall i}} \bigotimes_K \binom{N_K}{\boldsymbol{n_K}}^{1/2} \\
&\qquad\qquad\qquad\qquad \left[\prod_i \alpha_{Ki}^{n_{Ki}} \right] |\boldsymbol{n_K}\rangle_K,
\end{aligned}
\tag{D.116}
$$

which is of the same form as (D.114), up to the coefficients $\binom{N}{N_A,\ldots,N_Z}^{1/2} \prod_{K,i} \alpha_{Ki}^{n_{Ki}}$.