Dissertation

submitted to the

Combined Faculties for the Natural Sciences and for Mathematics

of the

Ruperto-Carola University of Heidelberg

for the degree of

Doctor of Natural Sciences

put forward by

Tobias Haas

born in Wermelskirchen, Germany

Oral examination: 26.04.2022

AN ENTROPIC PERSPECTIVE

ON

EQUILIBRIUM, UNCERTAINTY AND ENTANGLEMENT

Dedicated to those who play against all odds.
And to Bella, in $\sum_{n=1}^{\infty} \frac{1}{n}$ gratefulness.

"Natural science does not simply describe and explain nature;
it is part of the interplay between nature and ourselves;
it describes nature as exposed to our nature of questioning."

— Werner Heisenberg [1]

ABSTRACT

Entropy may be regarded as the physical quantity with the most facets. It is successfully used to describe aspects of the three intriguing phenomena: equilibrium, uncertainty and entanglement.

Despite its central meaning, formulating physical laws in terms of entropy can become unfavorable, for which we give three examples: the lack of well-defined continuum limits often prevents universal descriptions for discrete, continuous and infinite-dimensional degrees of freedom. The entropy of a subregion in a quantum field theory exhibits an ultraviolet divergence, which can not be renormalized. Entropies of marginal distributions do not capture the full information about a global distribution.

The first two problems can be traced back to the entropy being an absolute measure of missing information. To overcome these, we propose a more regular use of relative entropy in situations where entropy shows its flaws. Relative entropy allows us to unify entropic descriptions or to extend them into new regimes of validity.

In this thesis, we use relative entropy to formulate a new principle of inference, to develop thermodynamics in terms of model states, to derive divergence-free second law-like inequalities for relativistic fluids, to unify entropic uncertainty relations for discrete and continuous variables and to deduce the first entropic uncertainty relation for a quantum field.

The third problem becomes relevant in the context of entanglement witnessing for continuous variable systems. In contrast to standard separability criteria, which are based on measuring two observables separately, we start from a phase space representation of the quantum state. We find a perfect witness for pure state entanglement and derive entropic and even more general separability criteria, allowing us to certify entanglement in undetected regions.

ZUSAMMENFASSUNG

Die Entropie kann als die facettenreichste physikalische Größe angesehen werden. Man verwendet sie, um Aspekte von drei faszinierenden Phänomenen zu beschreiben: Gleichgewicht, Unschärfe und Verschränkung.

Trotz ihrer zentralen Bedeutung kann es ungünstig sein, physikalische Gesetze mit Entropie zu formulieren, wofür wir drei Beispiele anführen: Das Fehlen wohldefinierter Kontinuumslimites verhindert oft universelle Beschreibungen für diskrete, kontinuierliche und unendlich dimensionale Freiheitsgrade. Die Entropie einer Subregion in einer Quantenfeldtheorie weist eine ultraviolette Divergenz auf, die sich nicht renormieren lässt. Entropien von Marginalverteilungen erfassen nicht die vollständige Information über eine globale Verteilung.

Die ersten beiden Probleme lassen sich darauf zurückführen, dass die Entropie ein absolutes Maß für fehlende Informationen ist. Um diese zu überwinden, schlagen wir eine regelmäßigere Verwendung der relativen Entropie in Situationen vor, in denen die Entropie ihre Schwachstellen zeigt. Die relative Entropie ermöglicht es uns, entropische Beschreibungen zu vereinheitlichen oder sie auf neue Geltungsbereiche zu erweitern.

In dieser Arbeit verwenden wir die relative Entropie, um ein neues Inferenzprinzip zu formulieren, die Thermodynamik mit Modellzustände zu entwickeln, divergenzfreie zweite Hauptsatz-ähnliche Ungleichungen für relativistische Fluide abzuleiten, entropische Unschärferelationen für diskrete und kontinuierliche Variablen zu vereinheitlichen und die erste entropische Unschärferelation für ein Quantenfeld herzuleiten.

Das dritte Problem wird im Zusammenhang mit dem Nachweis von Verschränkung für Systeme mit kontinuierlichen Variablen relevant. Im Gegensatz zu standardmäßigen Separabilitätskriterien, die auf der getrennten Messung zweier Observablen beruhen, gehen wir von einer Phasenraumdarstellung des Quantenzustands aus. Wir finden ein perfektes Kriterium für die Verschränkung reiner Zustände und leiten entropische und noch allgemeinere Separabilitätskriterien ab, die es uns ermöglichen, Verschränkung in unentdeckten Regionen nachzuweisen.

PUBLICATIONS

Most ideas presented in this thesis are contained in a paper in preparation [A] and the six publications [B–G]. Detailed statements of contribution can be found at the beginning of every chapter. Besides, I was involved in the works [H–J] concerned with analog gravity models in Bose-Einstein condensates, which are not discussed in this thesis and will be available online soon.

[A] S. Floerchinger, M. Gärttner, **T. Haas**, and O. R. Stockdale, *General entanglement criteria in phase space from concave functions*, In preparation, 2022.

[B] S. Floerchinger, **T. Haas**, and M. Schröfl, "Relative entropic uncertainty relation for scalar quantum fields," Accepted by SciPost Physics, 1–13 (2022).

[C] S. Floerchinger, M. Gärttner, **T. Haas**, and O. R. Stockdale, "Entropic entanglement criteria in phase space," Phys. Rev. A **105**, 012409 (2022).

[D] S. Floerchinger, **T. Haas**, and H. Müller-Groeling, "Wehrl entropy, entropic uncertainty relations, and entanglement," Phys. Rev. A **103**, 062222 (2021).

[E] S. Floerchinger, **T. Haas**, and B. Hoeber, "Relative entropic uncertainty relation," Phys. Rev. A **103**, 062209 (2021).

[F] S. Floerchinger and **T. Haas**, "Thermodynamics from relative entropy," Phys. Rev. E **102**, 052117 (2020).

[G] N. Dowling, S. Floerchinger, and **T. Haas**, "Second law of thermodynamics for relativistic fluids formulated with relative entropy," Phys. Rev. D **102**, 105002 (2020).

[H] S. Floerchinger, **T. Haas**, Á. Parra-López, N. Sánchez-Kuntz, and M. Tolosa-Simeón, *Scalar quantum fields in cosmologies with $2+1$ spacetime dimensions*, In preparation, 2022.

[I] S. Floerchinger, **T. Haas**, M. Hans, E. Kath, N. Liebster, M. K. Oberthaler, Á. Parra-López, N. Sánchez-Kuntz, M. Sparn, H. Strobel, M. Tolosa-Simeón, and C. Viermann, *Curved and expanding space-time geometries in Bose-Einstein condensates*, In preparation, 2022.

[J] S. Floerchinger, **T. Haas**, M. Hans, E. Kath, N. Liebster, M. K. Oberthaler, Á. Parra-López, N. Sánchez-Kuntz, M. Sparn, H. Strobel, M. Tolosa-Simeón, and C. Viermann, *Engineering of curved space-time geometries for quantum fields*, In preparation, 2022.

NOTATION

Throughout this thesis, we work in terms of natural units $\hbar = c = k_B = 1$ and drop operator hats. Instead, operators are denoted by bold symbols, e.g. $\boldsymbol{X}$, while vectors are equipped with a vector symbol, e.g. $\vec{x}$, and expectation values or classical variables are written in normal font, e.g. X. The number of spacetime dimensions is labeled with d, while for the Hilbert space dimension we use $D = \dim \mathcal{H}$. Vacuum quantities are denoted with a bar, e.g. $\bar{\rho}$. A composite system is called 12 with the subsystems 1 and 2. We generically use the symbol S for every entropy and specify it further by its argument. For example, $S(\boldsymbol{\rho})$ denotes the von Neumann entropy of the quantum state $\boldsymbol{\rho}$. However, what is denoted as argument is sometimes not the argument of any entropy functional in a strict sense. This applies in particular to a conditional entropy $S(\boldsymbol{\rho}_1|\boldsymbol{\rho}_2)$ and a mutual information $I(\boldsymbol{\rho}_1 : \boldsymbol{\rho}_2)$. Hence, entropic quantities are particularized based on their meaning.

ACRONYMS

QM   Quantum Mechanics

QFT   Quantum Field Theory

UV   Ultraviolet

EUR   Entropic Uncertainty Relation

MU   Maassen-Uffink

BBM   Białynicki-Birula-Mycielski

FL   Frank-Lieb

REUR   Relative Entropic Uncertainty Relation

WL   Wehrl-Lieb

CPTP   Completely Positive Trace-Preserving

PTP   Positive Trace-Preserving

POVM   Positive Operator-Valued Measure

PVM   Projection-Valued Measure

LOCC   Local Operations and Classical Communication

REoE   Relative Entropy of Entanglement

PPT   Positive Partial Transpose

EPR   Einstein-Podolsky-Rosen

MGVT   Mancini-Giovannetti-Vitali-Tombesi

LIST OF FIGURES

CONTENTS

1 | INTRODUCTION

> *"[...] every physical quantity, every it,*
> *derives its ultimate significance from bits,*
> *binary yes-or-no indications,*
> *a conclusion which we epitomize in the phrase,*
> *it from bit."*
>
> — John Archibald Wheeler [2]

MOTIVATION Modern physics provides a fascinating richness of ideas and theories, including for example Quantum Mechanics (QM), Quantum Field Theory (QFT) and General Relativity. What these theories have in common is that they shifted well-established paradigms and were astonishingly successful in setting new standards on how precise experimental outcomes can be predicted.

However, at first sight, they are concerned with different phenomena on well separated scales and make use of rather dissimilar mathematical methods and concepts. This not only raises the desire for unification, but more importantly leads to the pressing question *what is fundamental*?

Some physicists may answer *quantum fields*, others may say *spacetime*. Many years after Shannon published his groundbreaking work on classical information theory [3, 4], *information* has become a valid alternative. Most importantly, Shannon gave the abstract concept of information a rigorous framework by establishing *entropy* as *the* measure for the missing information about a probability distribution.

At that time, entropy played an important role in physics already, most prominently in the contexts of statistical physics [5–7], quantum physics [8] and Szilard's treatment of Maxwell's demon [9]. However, it was Shannon's information theoretic approach which related entropy and information, showing that information is physical [10].

In the last three decades, the information theoretic viewpoint became a lot more common among physicists, especially in the context of quantum physics [11–13]. Nowadays physical laws are often expressed in terms of information theoretic quantities and quantum systems are regularly analyzed with information theoretic methods. Timely examples include quantum computing [14, 15] or the black hole information paradox [16–18].

However, some entropic descriptions are not satisfactory. In this thesis, we develop several methods to overcome fundamental problems of the entropy functional or to sharpen entropic statements. To that

end, we make use of two underestimated concepts: relative entropy [19–21] and entropy in quantum mechanical phase space [22, 23].

The relative entropy is a measure for the *distinguishability* of two probability distributions. As a *relative* measure of information, it shows rather universal properties. Especially, it often remains finite and keeps its properties when the entropy does not, for example when taking limits. Further, the entropy in phase space allows for a more complete description as opposed to considering the entropies of position and momentum separately.

We will apply our methods to describe the three phenomena equilibrium, uncertainty and entanglement. The problems or disadvantages of established entropic descriptions thereof and our proposals to resolve these issues are discussed for each phenomenon separately.

EQUILIBRIUM   Classical distributions and quantum states describing equilibrium situations can be derived from Jaynes' maximum entropy principle [24–27]. More precisely, the three thermodynamic ensembles microcanonical, canonical and grand canonical follow when maximizing entropy and assuming proper normalization and fixed expectation values for energy and particle number, respectively.

Although the maximum entropy principle was designed for this task, its applicability goes far beyond statistical physics. It is a fundamental principle of inference and as such predicts the optimal prior distribution or quantum state when only little information (often in terms of expectation values of macroscopic quantities) is given. In simple words, it expresses the intuition that the best guess should have least information content [27].

In Chapter 5, we formulate an alternative: the principle of minimum expected relative entropy for classical distributions (5.20) and quantum states (5.39). It postulates that the optimal prior is least distinguishable from other distributions on average and corresponds to the central point on the manifold of allowed distributions. Also, it closes a gap to the well-established principle of minimum discrimination information [20]. Although the latter is formulated in terms of relative entropy as well, it rather aims at updating a given prior distribution when additional information becomes available.

With the principle of minimum expected relative entropy at hand we reformulate thermodynamics in terms of relative entropy in Chapter 6, for which our motivation is twofold.

First, we want to extend the concepts of thermodynamics such as temperature and chemical potential beyond the situation where a system is in equilibrium. We propose a formulation in terms of thermodynamic reference states, which allows us to assign temperature and chemical potential to arbitrary states, see e.g. (6.28). Further, relative entropy captures thermal fluctuations (6.12) and encodes naturally

the second law of thermodynamics (6.35) as a special case of a central theorem about how quantum information can be processed.

Second, we want to understand better how thermodynamic concepts can be applied locally in the context of fluid dynamics and QFT. Here, entropy comes with two severe drawbacks. In fluid dynamics, one typically works in terms of entropy current densities, which are only well-defined in equilibrium. In addition, if one tries to define such an entropy from a more fundamental QFT, one has to deal with unavoidable Ultraviolet (UV) divergences due to entanglement. In contrast, a local relative entropy is free of such divergences and allows one to work with thermodynamic model states outside of equilibrium.

We apply the ideas gathered in Chapter 6 to fluids and quantum fields in Chapter 7. We establish a heuristic notion of local equilibrium with relative entropy in (7.19) and derive divergence-free second law-like inequalities in terms of relative entropy (7.37) from first principles.

UNCERTAINTY   For a long time the uncertainty principle was expressed through variances by default. However, such descriptions are unsatisfactory from an information theoretic perspective. The variance represents only the second moment of a measured probability distribution and lacks information about the full distribution.

Nowadays, the uncertainty principle is conveniently formulated in terms of Entropic Uncertainty Relations (EURs) [28, 29]. The three most well-known relations are the Maassen-Uffink (MU) relation [30, 31] for discrete variables (e.g. spin observables) and the Białynicki-Birula-Mycielski (BBM) [32] and Frank-Lieb (FL) [33] relations for continuous variables (e.g. position and momentum). Interestingly, the BBM relation is strictly stronger and implies the well-known relation by Heisenberg and Kennard, showing the supremacy of an entropic formulation.

In the field of EURs, entropy demonstrates its longest known flaw. Overlooked by Shannon and noticed by Jaynes, the entropy functional for discrete probabilities does not have a well-defined analog for probability density functions [26]. More precisely, the so-called differential entropy for probability densities (also defined by Shannon in [4]) does *not* correspond to the continuum limit of Shannon's discrete entropy. Instead, an infinite negative constant has to be added, breaking fundamental properties such as non-negativity.

Hence, it is not surprising that the two types of observables are described by different EURs. So far, only the MU and the FL relations have been unified, based on a formal measure-theoretic argument [33]. We propose a more direct unification in terms of relative entropy, the Relative Entropic Uncertainty Relation (REUR) (8.34), which we formulate in Chapter 8.

This approach demonstrates its advantages when considering quantum fields. In Chapter 9, we show that the field theory limit leads to divergences for the entropy associated with the probability density

functional of field configurations. These divergences are generic and well-known, for example from the vacuum energy, rendering every EUR ill-defined in the field theory limit. In contrast, relative entropy does not exhibit these divergences, allowing us to derive the first well-defined REUR for a scalar quantum field (9.36).

Also, we analyze EURs in canonical phase space by comparing relations based on marginal distributions with the Wehrl-Lieb (WL) relation [23, 34] formulated for a full phase space distribution in Chapter 10. In particular, we investigate which relation is closest to equality for a generic set of states. This is not only interesting by itself but is also very relevant for entanglement detection.

ENTANGLEMENT   Entanglement is the defining feature of quantum physics when considering one (or more) partitions of a system. Entanglement can cause the global entropy to be smaller than any local entropy, which is impossible for classical systems [35]. It is nowadays believed that entanglement plays an important role for several phenomena, including for example the thermalization of isolated quantum systems [36–38], black hole evaporation [39–41] and quantum supremacy [14].

However, it is rather hard to determine whether a given state is entangled or not, which is the so-called separability problem [42]. This problem remains unsolved except for a few special cases, for instance, Hilbert spaces of small dimensions and Gaussian states. Continuous variable systems with an infinite-dimensional Hilbert space prove to be particularly challenging for certifying entanglement.

Typically, one formulates conditions which are fulfilled by all separable states, so-called separability criteria, whose violation shows entanglement. For continuous variables, criteria were derived for the second moments of position and momentum measurements [43–46]. Also entropic criteria, which are stronger than the latter especially in the non-Gaussian regime, have been found [47–51].

These criteria have two things in common: First, they all make use of an uncertainty relation and second, they are all based on measuring position and momentum separately. Motivated by our findings in Chapter 10, we formulate several separability criteria for a *full* phase space distribution, pushing the boundaries of entanglement witnessing in continuous variable systems.

In Chapter 11, we show that a mutual information of a full phase space distribution (11.11) serves as a perfect witness for pure state entanglement. Then, we derive entropic criteria for mixed states (12.15) and show their outperfomance over other entropic criteria in Chapter 12. Finally, we deduce a more general set of criteria (13.10) in Chapter 13 for concave functions. Their experimental implementation is outlined by applying them to discretized distributions (13.26).

Part I

A PRIMER ON ENTROPY

We begin with introducing the three phenomena equilibrium, uncertainty and entanglement and the information theoretic methods required for a modern description thereof. In particular, we discuss the role of entropy within classical information (Chapter 2), quantum information (Chapter 3) and quantum fields (Chapter 4). An emphasis is put on how entropic quantities can be used to describe the physical phenomena we are concerned with.

2 | CLASSICAL INFORMATION

In simple words, classical information theory is concerned with probability distributions and how the information they contain about the underlying random variable can be quantified, stored and communicated. The theory goes back to early ideas by Nyquist [52, 53], Hartley [54], the groundbreaking papers of Shannon [3, 4] (see also [55]) and the follow-up work by Kullback and Leibler [19] (see also [20]) and Jaynes [24, 25] (see also [26, 27]).

Despite the success of QM, classical information theory is still of great importance to modern-day research. Any measurement, independent of whether the system of interest is well-described by a classical or a quantum theory, produces classical probability distributions, which have to be analyzed with the tools of classical information theory.

A rather significant distinction has to be made between discrete and continuous variables, i.e. between probabilities and probability densities. In the following, we include both cases and point out important differences. We follow standard literature on this subject [11, 12, 56].

2.1 PROBABILITY DISTRIBUTIONS AND DENSITIES

RANDOM VARIABLES Let us consider a random variable X which can take values x in some set $\mathcal{X}$, called alphabet, i.e. $x \in \mathcal{X}$. Whether the random variable X is called *discrete* or *continuous* depends on the cardinality of the alphabet $\mathcal{X}$. For discrete variables, $\mathcal{X}$ is finite or at least countable, whereas continuous variables are characterized by $\mathcal{X}$ being uncountable, such that

$$\mathcal{X} = \begin{cases} \{e_i\}_{i \in I} & \text{with} \quad I \subseteq \mathbb{N} & \text{for discrete } X, \\ (a,b) & \text{with} \quad a,b \in \mathbb{R} & \text{for continuous } X. \end{cases} \tag{2.1}$$

PROBABILITY In the case of discrete variables, one can associate a real number to every event $x \in \mathcal{X}$, which we call *probability $p(x)$*. Instead, for continuous variables, we work with a real-valued function $f : (a,b) \to \mathbb{R}$, which we refer to as the *probability density $f(x)$*. In both cases, we speak of a *distribution*.

The probability $p(x)$ and the probability density $f(x)$ can be fully characterized by the three Kolmogorov axioms [58], which require that they are

"[...] the theory of probabilities is at bottom only common sense reduced to calculus." - Pierre-Simon Laplace [57]

1. non-negative

$$p(x), f(x) \geq 0 \text{ for all } x \in \mathcal{X}, \tag{2.2}$$

2. normalized to unity

$$\sum_{x \in \mathcal{X}} p(x) = \int_{x \in \mathcal{X}} dx \, f(x) = 1, \tag{2.3}$$

3. additive for mutually exclusive events $x_i \in \mathcal{X}_i$ of countably many disjoint sets $\mathcal{X}_i \subseteq \mathcal{X}$

$$p\left(\cup_i x_i\right) = \sum_i p(x_i). \tag{2.4}$$

The last axiom holds equivalently for continuous variables. This can be seen after associating a probability for measuring a value in some finite interval, i.e. $x \in \mathcal{X}_j \subseteq \mathcal{X}$ with $\mathcal{X}_i$ being an uncountable subset of the full alphabet $\mathcal{X}$, by integrating over all elements in $\mathcal{X}_j$, such that

$$p(x \in \mathcal{X}_j) = \int_{x \in \mathcal{X}_j} dx \, f(x). \tag{2.5}$$

Furthermore, the three axioms imply many well-known basic rules of probability theory, for example that probabilities are bounded from above, i.e. $p(x) \leq 1$ for all $x \in \mathcal{X}$ (while probability densities remain unbounded) or the sum rule for non mutually exclusive events.

EXAMPLE: DISCRETE AND CONTINUOUS DISTRIBUTIONS To illustrate Kolmogorov's axioms and the conceptual differences between the two cases of variables, we consider (a) probabilities of an unfair die with $\mathcal{X} = \{1, 2, 3, 4, 5, 6\}$ (Figure 2.1a) and (b) a Gaussian probability density $f(x)$ of variance $\sigma^2 = 1$ with $\mathcal{X} = \mathbb{R}$ (Figure 2.1b).

(a) Discrete probabilities (b) Continuous probability density
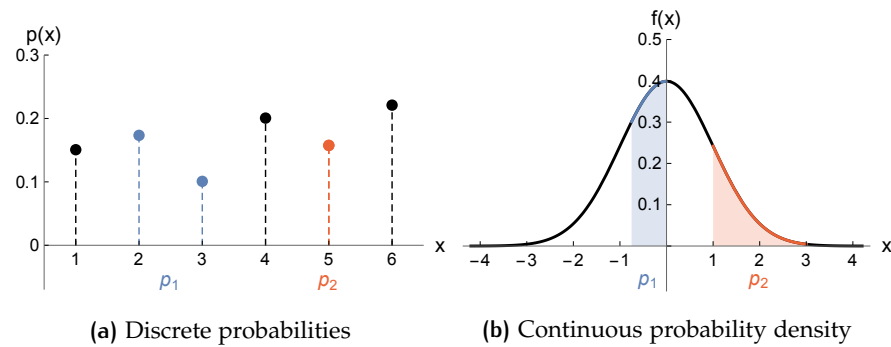
Figure 2.1: Comparison of (a) discrete probabilities and (b) a continuous probability density. The non-negativity and normalization axioms can be read off, while the colored events/regions indicate probabilities $p_1 \approx 0.273$ and $p_2 \approx 0.157$ of mutually exclusive sets computed by using the additivity axiom.

JOINT AND CONDITIONAL PROBABILITY The previous discussion can be extended to *two* (or more) random variables. Here, we additionally consider a second random variable Y with analogous notation as above. Then, we can associate a joint probability $p_{12}(x, y)$ or a joint probability density $f_{12}(x, y)$, respectively, to the joint measurement of X and Y, which also fulfill Kolmogorov's axioms.

Individual distributions $p_1(x)$ or $f_1(x)$ (analogously for the other variable Y) can be computed from marginalization

$$p_1(x) = \sum_{y \in \mathcal{Y}} p_{12}(x, y), \quad f_1(x) = \int_{y \in \mathcal{Y}} \mathrm{d}y \, f_{12}(x, y). \tag{2.6}$$

Moreover, if one of the two observables is measured, e.g. measuring Y gives y, one can associate a conditional probability $p_1(x|y)$ or conditional probability density $f_1(x|y)$ to the remaining random variable X, which are defined via

$$p_1(x|y) = \frac{p_{12}(x, y)}{p_2(y)}, \quad f_1(x|y) = \frac{f_{12}(x, y)}{f_2(y)}. \tag{2.7}$$

Note that in the latter case it may be necessary to require strict positivity of the marginal density $f(y) > 0$.

If and only if the two random variables X and Y are uncorrelated, the joint distributions factorize, i.e.

$$p_{12}(x, y) = p_1(x) \times p_2(y), \quad f_{12}(x, y) = f_1(x) \times f_2,(y) \tag{2.8}$$

and the conditional distributions reduce to the respective marginals

$$p_1(x|y) = p_1(x), \quad f_1(x|y) = f_1(x). \tag{2.9}$$

2.2 CLASSICAL ENTROPIES

SHANNON AND DIFFERENTIAL ENTROPY Arguably one of the most important contributions to information theory came from Shannon in the invention of a reasonable measure for the missing information about a distribution, which is the entropy. Although entropy is *the* fundamental quantity in information theory, it can *not* be consistently defined for discrete and continuous random variables [26]. This has major consequences on how physical laws have to be formulated, which will be one of our central topics.

Let us begin with the discrete case, where the entropy is named after Shannon himself. The entropy $S(p)$ of a discrete probability distribution $p(x)$ can be constructed axiomatically (up to a multiplicative constant) [3], but we rather follow an intuitive approach based on the notion of surprise [59]. To any event x one can associate a surprise $-\ln p(x)$, which is motivated by the surprise being small when

the event x is likely and additive for independent events. Then, the *Shannon entropy* is defined as the average surprise

$$S(p) = -\sum_{x \in \mathcal{X}} p(x) \ln p(x). \tag{2.10}$$

It serves as a measure for the missing information or uncertainty about $p(x)$. Also, it is non-negative $S(p) \geq 0$ (zero if and only if one event x is certain) and attains a global maximum $S(p) = \ln N$ for the uniform distribution $p(x) = \frac{1}{N}$ of a finite alphabet $|\mathcal{X}| = N$.

For continuous variables, the *differential entropy* is defined as

$$S(f) = -\int_{x \in \mathcal{X}} \mathrm{d}x \, f(x) \ln f(x). \tag{2.11}$$

It can be interpreted as a measure for the localization of a probability density $f(x)$ and, in contrast to the Shannon entropy $S(p)$, the differential entropy $S(f)$ can become negative if a distribution is highly localized. Also, it is in general not invariant under a reparameterization of the underlying random variable $x \rightarrow x'(x)$. Hence, an interpretation as a measure for missing information similarly to the Shannon entropy is not possible. In fact, the two definitions (2.10) and (2.11) are related via a negative infinite offset

$$S(f) = \lim_{\Delta x \to 0} \left(S(p) + \ln \Delta x \right), \tag{2.12}$$

where Δx is the width of a bin used to discretize $f(x)$. However, both entropies are concave, i.e. for a mixture of distributions $p(x) = \sum_i \lambda(i) p_i(x)$ or $f(x) = \sum_i \lambda(i) f_i(x)$ with $\lambda(i)$ being a discrete probability distribution itself, we obtain

$$S(p) \geq \sum_i \lambda(i) S(p_i), \quad S(f) \geq \sum_i \lambda(i) S(f_i). \tag{2.13}$$

EXAMPLE: ENTROPIES OF A COIN AND A GAUSSIAN We consider the entropies of (a) a coin with alphabet $\mathcal{X} = \{\text{heads}, \text{tails}\}$ and a probability p_1 for heads (Figure 2.2a) and (b) a centered Gaussian probability density $f(x)$ of standard deviation σ (Figure 2.2b).

RELATIVE ENTROPY Another fundamental quantity in information theory is the *relative entropy*, also known as Kullback-Leibler divergence [19, 20]. It turns out that the properties of the relative entropy are universal in many contexts, which we will utilize often in this thesis. Most importantly, it can consistently be defined for discrete as well as continuous variables, i.e. it has a well-defined continuum limit. Namely, for two probability distributions $p(x)$ and $\tilde{p}(x)$ it reads

$$S(p \| \tilde{p}) = \sum_{x \in \mathcal{X}} p(x) \left(\ln p(x) - \ln \tilde{p}(x) \right), \tag{2.14}$$

while for two probability densities $f(x)$ and $\tilde{f}(x)$ it is

$$S(f \| \tilde{f}) = \int_{x \in \mathcal{X}} \mathrm{d}x \, f(x) \left(\ln f(x) - \ln \tilde{f}(x) \right). \tag{2.15}$$

(a) Shannon entropy of a coin   **(b)** Differential entropy of a Gaussian
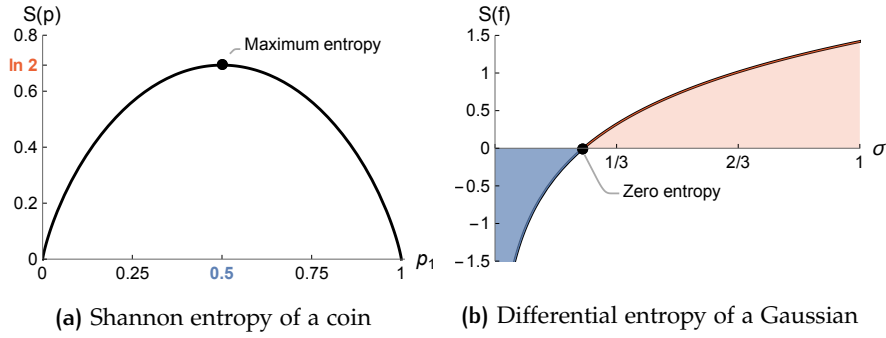
Figure 2.2: Comparison of (a) Shannon and (b) differential entropy. While the former is non-negative and takes its maximum value $\ln 2$ for the uniform distribution $p_1 = \frac{1}{2}$, the latter becomes negative for $\sigma < \frac{1}{\sqrt{2\pi e}}$ and grows monotonically with the broadness of the underlying distribution. Both functions are concave.

By Jensen's inequality, the relative entropy is always non-negative $S(p\|\tilde{p}), S(f\|\tilde{f}) \geq 0$ and zero if and only if the two distributions agree. Also, it is jointly convex in its two arguments. Hence, it can be interpreted as a measure for the *distinguishability* of the two distributions $p(x)$ and $\tilde{p}(x)$ (or $f(x)$ and $\tilde{f}(x)$). Nevertheless, it does not serve as a distance measure on the space of probability distributions, as it is neither symmetric nor fulfills a triangle inequality.

For our purposes, the second argument $\tilde{p}(x)$ (or $\tilde{f}(x)$) can be interpreted as a *model* or *reference* distribution for the true distribution given in the first argument $p(x)$ (or $f(x)$). Then, relative entropy measures the uncertainty deficit about the true distribution given the model.

For the relative entropy to be finite, we have to require a support condition $\text{supp}\,[p(x)] \subseteq \text{supp}\,[\tilde{p}(x)]$ (or $\text{supp}\,[f(x)] \subseteq \text{supp}\,[\tilde{f}(x)]$). Otherwise, its value is set to $+\infty$ indicating that the model can be distinguished perfectly from the true distribution, as it predicts zero probability for some possible events.

Also, the relative entropy is additive for independent distributions, just as entropy itself. In contrast to entropy, relative entropy *is* invariant under reparameterizations $x \to x'(x)$, which we will utilize later.

JOINT ENTROPY, CONDITIONAL ENTROPY AND MUTUAL INFORMATION
As for probabilities, the notion of entropy can be extended to the case of more than one random variable. Following the discussion of joint and conditional probabilities for two random variables X and Y, we define the *joint* Shannon entropy as

$$S(p_{12}) = -\sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}} p_{12}(x,y) \ln p_{12}(x,y) \tag{2.16}$$

and the *conditional* Shannon entropy as

$$S(p_1|p_2) = -\sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}} p_{12}(x,y) \ln p_1(x|y). \tag{2.17}$$

Both definitions can be extended to the continuous variable case in a similar fashion as shown above for the differential entropy, with the differences to the discrete variable case pointed out earlier.

Assuming finiteness of all involved entropies we find

$$S(p_{12}) = S(p_1) + S(p_2|p_1) = S(p_2) + S(p_1|p_2), \tag{2.18}$$

revealing that the conditional Shannon entropy $S(p_1|p_2)$ can be interpreted as the remaining uncertainty encoded in $p_1(x|y)$ after observing the random variable Y. Non-negativity of conditional entropy implies

$$S(p_{12}) \geq S(p_1), S(p_2), \tag{2.19}$$

showing that joint uncertainty is never smaller than any marginal one.

At last, we define the *mutual information* as the relative entropy between the joint probability distribution of interest $p_{12}(x, y)$ and the uncorrelated model $p_1(x) \times p_2(y)$

$$I(p_1 : p_2) = S(p_{12}|p_1 \times p_2) = S(p_1) + S(p_2) - S(p_{12}). \tag{2.20}$$

This allows one to interpret the mutual information as a universal (i.e. for discrete and continuous variables) measure for correlations.

By Jensen's inequality, the mutual information is non-negative, which is equivalent to classical entropies being subadditive and that conditioning cannot increase entropy, i.e.

$$I(p_1 : p_2) \geq 0 \Leftrightarrow S(p_{12}) \leq S(p_1) + S(p_2) \Leftrightarrow S(p_1|p_2) \leq S(p_1). \tag{2.21}$$

All three relations carry over to the continuous variable case and equality holds if and only if X and Y are uncorrelated.

EXAMPLE: ENTROPIC VENN DIAGRAMS We illustrate the entropic quantities and their relations in terms of Venn diagrams in Figure 2.3 for discrete probability distributions $p_1(x)$ and $p_2(y)$ with unequal entropies $0 < S(p_1) < S(p_2)$.

The information content of a system can be entirely encoded in the correlations with another system.

$I(p_1{:}p_2) = 0$ $\quad I(p_1{:}p_2) > 0$ $\quad I(p_1{:}p_2) = S(p_1)$

$S(p_1|p_2)$ $\quad S(p_2|p_1)$ $\quad S(p_2|p_1)$

$S(p_{12}) = S(p_1) + S(p_2)$ $\quad S(p_{12}) < S(p_1) + S(p_2)$ $\quad S(p_{12}) = S(p_2)$

(a) No correlations $\quad$ **(b)** Some correlations $\quad$ **(c)** Full correlations
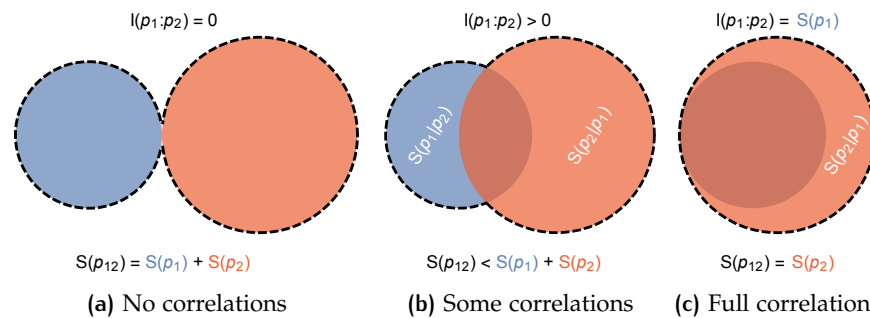
Figure 2.3: Venn diagrams for entropic quantities with increasing correlations from (a) to (c). Without correlations, entropy is additive. For some correlations, conditioning reduces marginal entropies. For full correlations, $p_2(y)$ contains all missing information.

3 | QUANTUM INFORMATION

In quantum information theory, classical information theoretic concepts are lifted to the quantum level. This means that information theoretic questions are directly addressed to the state of the quantum system of interest. In this sense, the classical probability distribution as a starting point is replaced by the quantum state.

In general, quantum information theory is concerned with similar questions as classical information theory. As quantum theory comes with a set of new physical phenomena, most notably entanglement and uncertainty, the aim of a quantum information theoretic treatment can also be to describe these phenomena as precisely as possible.

In the following, we give an introduction to the quantum information theoretic framework based on common literature [11–13, 59]. Thereupon, we describe the role of entropy in thermodynamics [24, 25, 61], measurement protocols and the quantum phase space [13, 62, 63], entropic uncertainty [28, 29, 64] and entanglement [42, 64, 65].

3.1 QUANTUM STATES AND QUANTUM CHANNELS

QUANTUM STATES   The *state* of a quantum system is fully described by the *density operator* ρ, which acts on the underlying Hilbert space $\mathcal{H}$ with dimension $D = \dim(\mathcal{H})$. Similar to Kolmogorov's construction of probabilities in Section 2.1, the density operator ρ can be characterized axiomatically. We denote the set of bounded linear operators from $\mathcal{H}$ to $\mathcal{H}$ as $\mathcal{B}(\mathcal{H})$. Then, the density operator $\rho \in \mathcal{B}(\mathcal{H})$ is defined to be

"[...] quantum mechanics is fundamentally a theory about the representation and manipulation of information, not a theory about the mechanics of nonclassical waves or particles." - Jeffrey Bub [66]

1. non-negative

$$\rho \geq 0, \quad \text{i.e.} \quad \langle \psi | \rho | \psi \rangle \geq 0 \text{ for all } |\psi\rangle \in \mathcal{H}, \tag{3.1}$$

2. normalized to unity

$$\mathrm{Tr}\{\rho\} = 1, \tag{3.2}$$

3. hermitian

$$\rho^\dagger = \rho. \tag{3.3}$$

By the spectral theorem, these axioms guarantee that there always exists a spectral decomposition in a complete orthonormal bases $\{|\psi\rangle\}_\psi$, such that

$$\rho = \sum_\psi p(\psi) |\psi\rangle \langle \psi|, \tag{3.4}$$

where $p(x)$ is a discrete probability distribution. In this basis, the matrix representation of ρ is diagonal.

Every density operator falls into one of two classes. Namely, ρ is either said to be *pure*, if and only if it is a projector $\rho^2 = \rho$, which is equivalent to $p(\psi) = 1$ for one particular ψ, such that $\rho = |\psi\rangle\langle\psi|$. If this is not the case, ρ is called *mixed*. This distinction can be subsumed with the two conditions $\text{Tr}\{\rho^2\} = 1$ or $\text{Tr}\{\rho^2\} < 1$, respectively. Also, (3.4) intuitively means that every quantum state ρ can be written as a convex combination of pure states (see also Figure 3.2).

QUANTUM CHANNELS Operations on a quantum state ρ are most generally modeled by *quantum channels*. Mathematically, a quantum channel is a linear map $\mathcal{N} : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$ acting on the state ρ, which is trace-preserving $\text{Tr}\{\mathcal{N}(\rho)\} = \text{Tr}\{\rho\} = 1$ and completely positive, i.e. $(\mathcal{N} \otimes \mathbb{1}_n)(\rho) \geq 0$ for all $n \in \mathbb{N}$. Consequently, it is referred to as a Completely Positive Trace-Preserving (CPTP) map. It can always be written as [67, 68]

$$\mathcal{N} : \rho \to \mathcal{N}(\rho) = \sum_\alpha A_\alpha \rho A_\alpha^\dagger, \tag{3.5}$$

where A_α are the so-called Kraus operators fulfilling the normalization condition $\sum_\alpha A_\alpha^\dagger A_\alpha = \mathbb{1}$. In the following, we discuss four important examples of quantum channels, which will play a major role in this thesis.

PARTIAL TRACE The notion of a quantum state can be extended to the case where a quantum system consists of two (or more) parts. This situation is often referred to as *bipartition* 12. Then, the system is described by a joint Hilbert space $\mathcal{H}_{12} = \mathcal{H}_1 \otimes \mathcal{H}_2$ with bases $\{|\psi\rangle\}_\psi$ and $\{|\phi\rangle\}_\phi$ for $\mathcal{H}_1$ and $\mathcal{H}_2$, respectively, to which we assign a so-called *global* state ρ_{12}. The *local* states of the individual systems 1 and 2 follow by tracing over complementary degrees of freedom, i.e.

$$\rho_1 = \text{Tr}_2\{\rho_{12}\} = \sum_\phi \langle\phi|\rho_{12}|\phi\rangle, \tag{3.6}$$

and analogously for ρ_2. Therein, the *partial trace* $\text{Tr}_i\{.\}$ with $i = 1, 2$ is our first example for a quantum channel.

TIME EVOLUTION OF A CLOSED SYSTEM For a closed quantum system, the time evolution of its quantum state ρ is described by the von Neumann equation

$$\partial_t \rho = -i[H, \rho], \tag{3.7}$$

where H denotes the Hamilton operator of the system and $[.,.]$ is the commutator. Note that we work in the Schrödinger picture where states are time-dependent $\rho = \rho(t)$ while observables are stationary. For pure states $\rho = |\psi\rangle\langle\psi|$, (3.7) reduces to the Schrödinger equation.

The von Neumann equation is formally solved by the self-contained expression

$$\rho(t_f) = U(t_f - t_i)\,\rho(t_i)\,U^\dagger(t_f - t_i), \tag{3.8}$$

where $t_f > t_i$ are two instances of time and $U(t_f - t_i)$ denotes the *unitary time evolution operator* from t_i to t_f, which reads

$$U(t_f - t_i) = e^{-iH(t_f - t_i)} \tag{3.9}$$

and fulfills $UU^\dagger = U^\dagger U = \mathbb{1}$. Hence, closed quantum systems evolve unitarily in time. Also, (3.8) shows that unitary time evolution is a quantum channel with the unitary time evolution operator (3.9) being the Kraus operator.

TIME EVOLUTION OF AN OPEN SYSTEM An open quantum system typically does not evolve unitarily, for example due to the exchange of energy with its environment. Exemplary, one may think of a subsystem of a closed system or a system which is in thermal contact with a heat bath. The time evolution of an open system is, depending on the initial conditions, described by a quantum channel $\mathcal{N}$ of the form (3.5), or at least by a Positive Trace-Preserving (PTP) map $\mathcal{M}$, which fulfills $\mathcal{M}(\rho) \geq 0$, but is not completely positive in general [69–71].

MEASUREMENTS In quantum information theory, *measurements* are mathematically described in the framework of Positive Operator-Valued Measures (POVMs). The measurement of an observable M can be represented as a collection of hermitian operators $\{M_m\}_m$, where m denotes discrete or continuous outcomes drawn from the alphabet $\mathcal{M}$ corresponding to the observable M. To those operators one can associate the non-negative POVM elements

$$E_m = M_m^\dagger M_m, \tag{3.10}$$

which are equipped with a resolution of the identity, i.e.

$$\mathbb{1} = \sum_m E_m, \quad \mathbb{1} = \int \mathrm{d}m\, E_m. \tag{3.11}$$

Then, the probability $p(m)$ or probability density $f(m)$ to obtain the measurement outcome m in the state ρ when measuring the observable M can be computed with the simple formulas

$$p(m) = \mathrm{Tr}\{\rho E_m\}, \quad f(m) = \mathrm{Tr}\{\rho E_m\}, \tag{3.12}$$

respectively.

A special case of a POVM is a Projection-Valued Measure (PVM), where all POVM elements E_m coincide with their corresponding measurement operators M_m. This is fulfilled if and only if all considered

elements are orthogonal projectors, i.e. $E_m = M_m^\dagger M_m = M_m$ for orthogonal measurement operators, i.e. $M_m M_{m'} = \delta_{mm'} M_m$ for discrete or $M_m M_{m'} = \delta(m - m') M_m$ for continuous outcomes m.

Two remarks regarding the measurements performed in experiments are in order. First, it is important to note that in experiments one always works with PVMs. However, Naimark's dilatation theorem ensures that every POVM can be represented as a PVM on an enlarged Hilbert space (see e.g. [72]). In practice, this requires coupling the system of interest to an ancilla system and performing a PVM on the joint system. Second, continuous measurements should be considered as an idealized situation. In actual experiments, measurement outcomes are always discrete as a consequence of finite resolution. We will further comment on this issue and its implications in Section 13.3.

EXAMPLE: QUANTUM CHANNELS The four discussed quantum channels are altogether depicted in Figure 3.1.
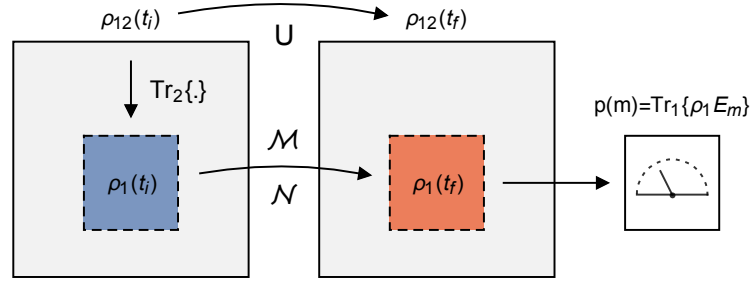
Figure 3.1: Illustration of the four discussed quantum channels: partial trace $\mathrm{Tr}_2\{.\}$, unitary time evolution of a closed system U, non-unitary time evolution of an open subsystem governed by a CPTP $\mathcal{N}$ or PTP map $\mathcal{M}$ and a measurement $\mathrm{Tr}_1\{\rho_1 E_m\}$.

3.2 QUANTUM ENTROPIES

VON NEUMANN ENTROPY The idea of a measure for the missing information about a probability distribution can straightforwardly be extended to the quantum level. To that end, the classical probability distribution is replaced by the full quantum state. Consequently, we define the *von Neumann entropy* as [8]

$$S(\rho) = -\mathrm{Tr}\{\rho \ln \rho\}. \tag{3.13}$$

Independent of whether the system of interest possesses discrete or continuous degrees of freedom, the von Neumann entropy is non-negative $S(\rho) \geq 0$ being zero if and only if the state is pure $\rho = |\psi\rangle\langle\psi|$. Hence, it can be thought of as a measure for the uncertainty about or the mixedness of the state ρ.

The von Neumann entropy $S(\rho)$ is concave in ρ, i.e. for a mixture $\rho = \sum_i \lambda(i)\rho_i$ with $\lambda(i)$ being a probability distribution, we obtain

$$S(\rho) \geq \sum_i \lambda(i)S(\rho_i). \tag{3.14}$$

Also, $S(\rho)$ equals the Shannon entropy of the mixture probabilities $p(\psi)$ if we evaluate it in the eigenbasis of ρ as given in (3.4),

$$S(\rho) = S(p), \tag{3.15}$$

implying that it takes its global maximum $S(\rho) = \ln D$ for the maximally mixed state $\rho = \sum_\psi \frac{1}{D} |\psi\rangle \langle\psi|$.

Another important property of the von Neumann entropy $S(\rho)$ is its invariance under unitary transformations U, i.e.

$$S(U\rho U^\dagger) = S(\rho). \tag{3.16}$$

Together with (3.8) this implies that the mixedness of a closed quantum system stays constant over time.

EXAMPLE: VON NEUMANN ENTROPY CURVES   We consider a three-state system whose Hilbert space $\mathcal{H}$ is spanned by $\{|0\rangle, |1\rangle, |2\rangle\}$ and a diagonal quantum state $\rho = \sum_{\psi=1}^3 p(\psi)|\psi\rangle \langle\psi|$. Curves of constant entropies $S(\rho)$ are shown in Figure 3.2.
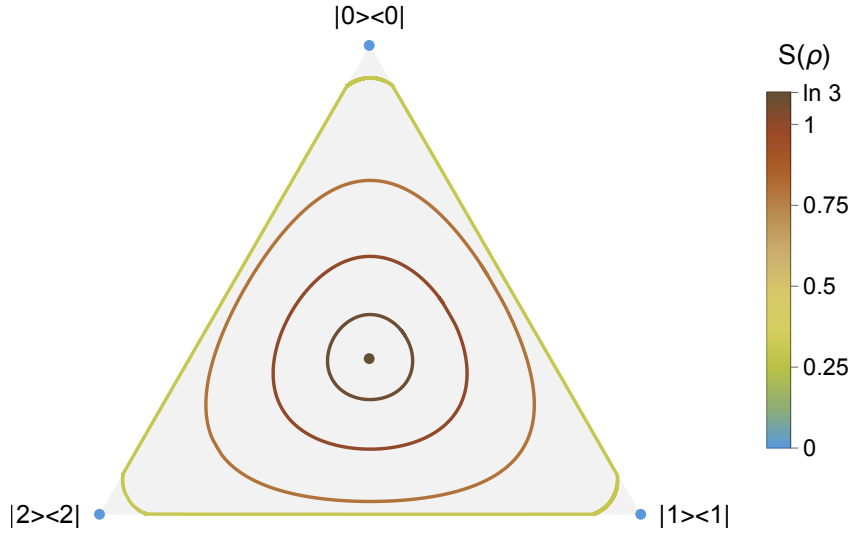
Figure 3.2: Curves of constant von Neumann entropy $S(\rho)$ for a three-level system on the simplex containing all states diagonal in the $\{|0\rangle, |1\rangle, |2\rangle\}$ basis. Convexity of the state space as well as increasing mixedness towards the center become visible.

QUANTUM RELATIVE ENTROPY   The *quantum relative entropy* is defined for two density operators ρ and $\tilde\rho$ as [21]

$$S(\rho\|\tilde\rho) = \mathrm{Tr}\{\rho(\ln\rho - \ln\tilde\rho)\}, \tag{3.17}$$

with the support condition $\text{supp}[\rho] \subseteq \text{supp}[\tilde{\rho}]$ understood. Analogous to its classical counterpart, it is a non-negative and jointly convex measure for the distinguishability of the two states ρ and $\tilde{\rho}$ being zero if and only if $\rho = \tilde{\rho}$.

Two states ρ and $\tilde{\rho}$ are typically harder to distinguish after a quantum operation.

The quantum relative entropy is used to formulate one of the most fundamental theorems in quantum information theory, which is referred to as *monotonicity of quantum relative entropy* under quantum channels $\mathcal{N}$ and reads

$$S(\rho\|\tilde{\rho}) \geq S(\mathcal{N}(\rho)\|\mathcal{N}(\tilde{\rho})) \tag{3.18}$$

with equality for unitary channels. This inequality has been proven also for PTP maps in [73] and was further strengthened in [74].

CONDITIONAL VON NEUMANN ENTROPY AND QUANTUM MUTUAL IN-FORMATION In case of a bipartition 12, we can associate a von Neumann entropy to the global state $S(\rho_{12})$ as well as to the two local states $S(\rho_1)$ and $S(\rho_2)$ in the sense of (3.13). As for the states themselves, these entropies are referred to as global and local entropies, respectively.

Then, one can define the *conditional* von Neumann entropy via

$$S(\rho_1|\rho_2) = S(\rho_{12}) - S(\rho_2) \tag{3.19}$$

as the remaining uncertainty about ρ_1 when knowing the state ρ_2, provided that all involved entropies are finite [75]. It is concave in ρ_{12}, but in opposition to the classical case, the quantum version of the conditional entropy can become negative,

$$S(\rho_{12}) \not\geq S(\rho_1), S(\rho_2), \tag{3.20}$$

which we will discuss in more detail in Section 3.6. Hence, quantum theory allows for more uncertainty in the parts of a system than in the system as a whole. The quantum analog of (2.19) is instead given by the triangle inequality

$$S(\rho_{12}) \geq |S(\rho_1) - S(\rho_2)|. \tag{3.21}$$

Note that if the global state is pure $S(\rho_{12}) = 0$, this implies that the two local von Neumann entropies agree $S(\rho_1) = S(\rho_2)$.

Also, we define a *quantum mutual information*

$$I(\rho_1 : \rho_2) = S(\rho_{12}\|\rho_1 \otimes \rho_2) = S(\rho_1) + S(\rho_2) - S(\rho_{12}), \tag{3.22}$$

which is still a non-negative measure for the (quantum and classical) correlations between the two subsystems 1 and 2.

In complete analogy to (2.21), non-negativity of quantum mutual information is equivalent to subadditivity of quantum entropies and that conditioning cannot increase quantum entropies, i.e.

$$I(\rho_1 : \rho_2) \geq 0 \Leftrightarrow S(\rho_{12}) \leq S(\rho_1) + S(\rho_2) \Leftrightarrow S(\rho_1|\rho_2) \leq S(\rho_1), \tag{3.23}$$

with equality if and only if the two subsystems are uncorrelated $\rho_{12} = \rho_1 \otimes \rho_2$.

3.3 MAXIMUM ENTROPY AND THERMODYNAMICS

THE MAXIMUM ENTROPY PRINCIPLE   So far, we have discussed the notions of probabilities, states and how to describe their information content with the methods of (quantum) information theory. However, there remains the (probably most) elementary question of probability theory: Given a physical system and some information about it, typically in terms of expectation values of some observables, which probability distribution or state should we assign to it? In other words, how can we find an unbiased *prior* or what is the *best guess*?

An early attempt to solve this problem was given by Laplace, which is referred to as the principle of indifference. It states that in absence of any further information, the best guess is given by the uniform distribution [27]. Although this approach appears to be natural, it lacks of applicability in many cases. For example, it cannot make any statements for continuous variables on unbounded intervals or when there is indeed access to side information.

With the invention of the entropy as a measure for the missing information of a distribution (or state), a more quantitative treatment of such a question became possible, which led Jaynes to propose his *maximum entropy principle* [24, 25].

According to this principle, the best guess is given by the distribution $p(x)$ or state ρ, which respects the available side information and maximizes the corresponding classical or quantum entropy $S(p)$ or $S(\rho)$, respectively, i.e.

$$\delta S(p) \overset{!}{=} 0, \quad \delta S(\rho) \overset{!}{=} 0, \tag{3.24}$$

with side constraints implemented via the method of Lagrange multipliers. Note that the normalization of probabilities or states is always the minimal constraint.

To provide some heuristic argument for the reasonability of this principle, let us consider the situation in which we have found two distributions $p(x)$ and $\tilde{p}(x)$, which both respect all available side information, and have entropies $S(p) < S(\tilde{p})$. As entropy is a measure for missing information, the latter implies that $p(x)$ contains additional information, which cannot be accessed by the experiment performed. Hence, we prefer $\tilde{p}(x)$ over $p(x)$ as our best guess should not include information we do not have access to.

Note also that in absence of side information, the maximum entropy principle reduces to the principle of indifference. Hence, it may be regarded as an extension of the latter.

INFORMATION THEORETIC APPROACH TO THERMODYNAMICS   Statistical physics may be regarded as the prime example for an application of the maximum entropy principle. In fact, the latter was mainly developed for justifying the distributions and states describing classical

"[...], the 'subjective' school of thought regards probabilities as expressions of human ignorance; the probability of an event is merely a formal expression of our expectation that the event will or did occur, based on whatever information is available. To the subjectivist, the purpose of probability theory is to help us in forming plausible conclusions in cases where there is not enough information available to lead to certain conclusions; thus detailed verification is not expected. The test of a good subjective probability distribution is does it correctly represent our state of knowledge as to the value of X?" - Edwin T. Jaynes [24]

and quantum equilibrium situations [24, 25]. Note that in the context of thermodynamics, entropies are typically multiplied with the Boltzmann constant k_B to match the definition of the thermodynamic Boltzmann entropy. As information theoretic entropies are only defined up to a multiplicative factor, one may easily adapt to this choice. For brevity, we keep natural units.

We illustrate the typical line of reasoning by making use of some simplifying assumptions. We consider a macroscopic quantum system in a finite volume V and a fixed particle number N, which is allowed to exchange energy with a heat bath of temperature T. Such a system may be described by a time-independent Hamiltonian $H \neq H(t)$. Now, we want to find the optimal prior ρ for this situation.

The Hamiltonian fulfills an eigenvalue equation

$$H \ket{n} = E(n) \ket{n}, \tag{3.25}$$

where $\{\ket{n}\}_n$ denotes the orthonormal basis of energy eigenstates with corresponding eigenvalues $E(n)$. For simplicity, we assume that these eigenvalues are not degenerate, i.e. $E(n) \neq E(m)$ for all $n \neq m$. Then, the von Neumann equation (3.7) implies that any stationary state $\rho \neq \rho(t)$ is diagonal in the energy eigenbasis

$$\rho = \sum_n p(n) \ket{n} \bra{n}, \tag{3.26}$$

where $p(n)$ is a discrete probability distribution. By the relation (3.15), the von Neumann entropy of such a state is equal to the Shannon entropy of $p(n)$, i.e. $S(\rho) = S(p)$, which can be interpreted as a classical limit and allows us to work with classical expressions in the following. Note that such a relation does not hold for degenerate eigenstates.

Furthermore, let us assume that we have access to the energy expectation value

$$E(p) = \mathrm{Tr}\{\rho H\} = \sum_n p(n) E(n). \tag{3.27}$$

Then, the maximum entropy principle states that the best guess for the distribution $p(n)$ (or the state ρ) is the result of maximizing the corresponding entropy $S(p)$ (or $S(\rho)$) under the side constraints of normalization and a given energy expectation value (3.27), implemented via Lagrangian multipliers λ and β, respectively. This optimization problem is solved by the canonical state

$$\rho_c = \sum_n p_c(n) \ket{n} \bra{n}, \tag{3.28}$$

with

$$p_c(n) = \frac{1}{Z_c} e^{-\beta E(n)}, \tag{3.29}$$

where $Z_c = \sum_n e^{-\beta E(n)}$ follows from normalization and can be identified with the canonical partition sum, from which all thermodynamic quantities can be derived. From the remaining side constraint (3.27) we find

$$E(\boldsymbol{\rho}_c) = -\frac{\partial}{\partial \beta} \ln Z_c, \tag{3.30}$$

such that the entropy becomes

$$S(\boldsymbol{\rho}_c) = \ln Z_c + \beta E(\boldsymbol{\rho}_c). \tag{3.31}$$

Following the standard definition of temperature

$$\frac{1}{T} \equiv \frac{\partial S(\boldsymbol{\rho}_c)}{\partial E(\boldsymbol{\rho}_c)}, \tag{3.32}$$

we can identify the second Lagrangian multiplier with the inverse temperature

$$\beta = \frac{1}{T}. \tag{3.33}$$

At last, we find that the free energy as the thermodynamic potential is related to the canonical partition sum Z_c via

$$F(\boldsymbol{\rho}_c) \equiv E(\boldsymbol{\rho}_c) - TS(\boldsymbol{\rho}_c) = -T \ln Z_c. \tag{3.34}$$

To summarize, the maximum entropy principle led us to a complete description in terms of the canonical ensemble $\boldsymbol{\rho}_c$.

We will reformulate this principle in terms of relative entropy and thereupon develop thermodynamics with relative entropy only in Chapter 5 and Chapter 6, respectively.

SECOND LAW OF THERMODYNAMICS Let us now come to the second law of thermodynamics, which may be considered as one of the most central observations about nature. In simple words, it expresses the tendency of nature to evolve towards states of higher entropy. In favor of our entropic perspective, we state the second law following Clausius' line of reasoning.

To that end, we reconsider the situation described before, i.e. a system of finite volume V with fixed particle number N coupled to a heat bath of temperature T, such that this system is well-described by the canonical state $\boldsymbol{\rho}_c$. Now, imagine that the system undergoes some process from time t_i to time $t_f > t_i$, in which it exchanges energy with the heat bath. Typically, one assumes processes to be quasi-stationary within the framework of thermodynamics, but as we will see later, such an assumption is not necessary to obtain second law-like inequalities. Then, *Clausius' inequality* states that the change in entropy $\Delta S(\boldsymbol{\rho}_c) = S_f(\boldsymbol{\rho}_c) - S_i(\boldsymbol{\rho}_c)$ is bounded from below

$$\Delta S(\boldsymbol{\rho}_c) \geq \beta \Delta E(\boldsymbol{\rho}_c), \tag{3.35}$$

"The law that entropy always increases - the second law of thermodynamics - holds, I think, the supreme position among the laws of Nature. If someone points out to you that your pet theory of the universe is in disagreement with Maxwell's equations - then so much the worse for Maxwell's equations. If it is found to be contradicted by observations - well, these experimentalists do bungle things sometimes. But if your theory is found to be against the second law of thermodynamics I can give you no hope; there is nothing for it but to collapse in deepest humiliation."
- Sir Arthur S. Eddington [76]

where $\Delta E(\boldsymbol{\rho}_c) = E_f(\boldsymbol{\rho}_c) - E_i(\boldsymbol{\rho}_c)$ denotes the change in energy in the system from time t_i to time t_f. The latter relation reduces to an equality if and only if the quasi-stationary processes are reversible.

In Chapter 6 we will derive second law-like inequalities from the monotonicity property of the quantum relative entropy, which we then generalize to a local QFT in the framework of relativistic fluids in Chapter 7.

BEYOND THE SCOPE Let us briefly mention other approaches to thermodynamics and the second law, which we do not further consider. Classical statistical physics may be developed starting from ergodic theory [61, 77], while in quantum statistical physics the eigenstate thermalization hypothesis has been put forward [78–82]. Alternatively, one may derive the thermodynamic ensembles from a general analysis of entanglement [36]. All these approaches also try to explain the process of *thermalization* and under which conditions it occurs.

Furthermore, considerable effort has been put into defining thermodynamic entropy axiomatically in an operational way, with the aim of extending its definition beyond the equilibrium case. Here, we refer to numerous works by Lieb and Yngvason [83–85] and Zanchini and Beretta [86–93]. These approaches are based on the concept of adiabatic accessibility as an ordering relation between quantum states and a set of axioms, from which also the second law can be derived.

3.4 MEASUREMENTS AND PHASE SPACE

CONTINUOUS VARIABLES: DIRECT MEASUREMENT In this introductory section we consider two degrees of freedom only. Generalizations to more degrees of freedom are discussed in Chapter 9 and Chapter 11.

We start with the case of continuous variables, which is characterized by an infinite-dimensional Hilbert space $D = \infty$. The typical pair of observables under consideration is *position X* and *momentum K*, whose algebra is given by the canonical commutation relation

$$[\boldsymbol{X}, \boldsymbol{K}] = i. \tag{3.36}$$

At first, we consider the so-called *homodyne detection protocol*, where the corresponding measurement operators are constructed from the eigenvalue equations

$$\boldsymbol{X}\ket{x} = x\ket{x}, \quad \boldsymbol{K}\ket{k} = k\ket{k}. \tag{3.37}$$

The eigenvectors are orthogonal but non-normalisable in the sense that

$$\braket{x|x'} = \delta(x - x'), \quad \braket{k|k'} = \delta(k - k'), \tag{3.38}$$

and are related via

$$\langle x|k\rangle = \frac{1}{\sqrt{2\pi}}e^{ikx}. \tag{3.39}$$

Then, we can define two PVMs from the corresponding sets of eigenvectors as $\{|x\rangle\langle x|\}_x$ and $\{|k\rangle\langle k|\}_k$. The probability densities follow from (3.12) as

$$f(x) = \langle x|\rho|x\rangle, \quad g(k) = \langle k|\rho|k\rangle, \tag{3.40}$$

and contain the full information about their respective underlying observable. However, they do not contain any information about the correlations between X and K, raising the question after other measurement schemes.

CANONICAL COHERENT STATES   Besides directly detecting position and momentum by measuring in their corresponding eigenbases, there is another possibility, for which we need to introduce *canonical coherent states*. We loosely follow the group-theoretic approach [94, 95] (see also [96–99]), which has the great advantage that the notion of coherent states can be generalized to other algebraic structures.

We start with defining creation and annihilation operators via

$$a^\dagger = \frac{1}{\sqrt{2}}(X - iK), \quad a = \frac{1}{\sqrt{2}}(X + iK), \tag{3.41}$$

which fulfill the commutation relations

$$[a, a^\dagger] = 1. \tag{3.42}$$

Together with the identity operator $\mathbb{1}$ and the number operator $N = a^\dagger a$, these four form Heisenberg-Weyl algebra H_4. We may define the vacuum state $|0\rangle$ as the extremal state in the sense that

$$a|0\rangle = 0. \tag{3.43}$$

Then, the isotropy group, i.e. the largest subgroup which leaves the vacuum invariant, is given by $U(1) \otimes U(1)$. The operators generating the canonical coherent states live in the corresponding coset space $H_4/U(1) \otimes U(1)$. A convenient unitary representation of coset space elements is given by the so-called *displacement operator*

$$D(\alpha) = e^{\alpha a^\dagger - \alpha^* a}, \tag{3.44}$$

which fulfills $D^\dagger(\alpha) = D^{-1}(\alpha) = D(-\alpha)$, with a complex-valued parameter α. The set of canonical coherent states follows from acting on the vacuum state $|0\rangle$ with the displacement operator

$$|\alpha\rangle = D(\alpha)|0\rangle. \tag{3.45}$$

Hence, canonical coherent states can geometrically be represented in the complex plane, which corresponds to the quantum phase space $\mathbb{R}^2$. Consequently, it is natural to work with the parameterization

$$\alpha = \frac{1}{\sqrt{2}} \left(x + ik \right). \tag{3.46}$$

Note that besides the group theoretic approach, one can also define canonical coherent states as eigenstates of the annihilation operator

$$\boldsymbol{a} \left| \alpha \right\rangle = \alpha \left| \alpha \right\rangle. \tag{3.47}$$

However, this definition does not allow for a straightforward generalization to other groups.

Let us proceed with listing three important properties of canonical coherent states. They are *not* orthogonal

$$|\langle \alpha | \beta \rangle|^2 = e^{-\frac{1}{2}\left((x-x')^2 + (k-k')^2 \right)}, \tag{3.48}$$

but nonetheless span an overcomplete basis

$$\mathbb{1} = \int \frac{\mathrm{d}x\,\mathrm{d}k}{2\pi} \left| \alpha \right\rangle \left\langle \alpha \right|, \tag{3.49}$$

and minimize several uncertainty relations (e.g. (3.80)).

CONTINUOUS VARIABLES: JOINT MEASUREMENT From (3.49) we can construct a POVM from pure coherent state projectors

$$\boldsymbol{E}_\alpha = \left| \boldsymbol{\alpha} \right\rangle \left\langle \boldsymbol{\alpha} \right|, \tag{3.50}$$

"The best thing one can do is to measure the probability of finding a particle in a state with minimal uncertainty centered around the classical values of momentum and position, i.e. in a coherent state." - Alfred Wehrl [23]

which is the so-called *heterodyne measurement*. It leads to the Husimi Q-distribution [100]

$$Q(x,k) = \mathrm{Tr}\{\boldsymbol{\rho} \boldsymbol{E}_\alpha\} = \left\langle \alpha | \boldsymbol{\rho} | \alpha \right\rangle, \tag{3.51}$$

which is a *quasi*-probability distribution in the quantum phase space and as such contains the information about both variables X and K *and* their correlations. It is normalized to unity with respect to the phase space measure

$$\int \frac{\mathrm{d}x\,\mathrm{d}k}{2\pi} Q(x,k) = \mathrm{Tr}\{\boldsymbol{\rho}\} = 1 \tag{3.52}$$

and bounded

$$0 \leq Q(x,k) \leq 1, \tag{3.53}$$

where the upper bound can heuristically be interpreted as a consequence of the uncertainty principle [101, 102]. However, it has to be understood as a quasi-probability distribution as coherent states are

not orthogonal and hence the Husimi Q-distribution violates the third axiom by Kolmogorov (2.4).

It is worth mentioning that the Husimi Q-distribution is related to the wider known Wigner W-distribution by a Weierstrass transform with respect to the vacuum, i.e.

$$Q(x,k) = \int \frac{\mathrm{d}x' \, \mathrm{d}k'}{2\pi} \, W(x',k') \, 2 \, e^{-(x-x')^2 - (k-k')^2}. \tag{3.54}$$

Hence, the Husimi Q-distribution can be interpreted as a broadened version of the Wigner W-distribution, which removes the negativities of the latter. Moreover, the probability densities $f(x)$ and $g(k)$ are in fact the marginals of the Wigner distribution

$$f(x) = \int \frac{\mathrm{d}k}{2\pi} \, W(x,k), \quad g(k) = \int \frac{\mathrm{d}x}{2\pi} \, W(x,k), \tag{3.55}$$

indicating that the marginals of the Husimi Q-distribution correspond to broadened versions of $f(x)$ and $g(k)$. For an overview of these and other quasi-probability distributions see [94, 102–105].

SYMPLECTIC STRUCTURE OF CANONICAL PHASE SPACE   It is often convenient to group the phase space operators X and K into a single vector-valued operator

$$\vec{\chi} = (X, K). \tag{3.56}$$

Then, the canonical commutation relations (3.36) can be rewritten as

$$[\chi_j, \chi_k] = i\Omega_{jk}, \tag{3.57}$$

where Ω is the *symplectic metric*

$$\Omega = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \tag{3.58}$$

which fulfills $\Omega^T = \Omega^{-1} = -\Omega$ and is an involution $\Omega^2 = -\mathbb{1}$.

GAUSSIAN STATES   For continuous variable systems we often deal with Gaussian states, i.e. states for which the Wigner W-distribution (or equivalently the Husimi Q-distribution) is of *Gaussian form*. Hence, these states are fully determined by their means

$$\chi_j = \mathrm{Tr}\{\rho\chi_j\} \tag{3.59}$$

and covariance matrix

$$\gamma_{jk} = \frac{1}{2}\mathrm{Tr}\{\rho(\chi_j\chi_k + \chi_k\chi_j)\} - \mathrm{Tr}\{\rho\chi_j\}\mathrm{Tr}\{\rho\chi_k\}, \tag{3.60}$$

which is by definition real, symmetric and positive semi-definite. It can also be written as

$$\gamma = \begin{pmatrix} \sigma_x^2 & \sigma_{xk} \\ \sigma_{xk} & \sigma_k^2 \end{pmatrix}, \tag{3.61}$$

with σ_x^2 and σ_k^2 denoting the variances of position and momentum measurements, respectively, while σ_{xk} is the covariance, highlighting that γ is the covariance matrix of the Wigner W-distribution. Via (3.54) we find for the covariance matrix V of the Husimi Q-distribution

$$V = \gamma + \frac{1}{2}\mathbb{1}. \tag{3.62}$$

Notably, Hudson's theorem states that a pure state is Gaussian if and only if its Wigner W-distribution is non-negative everywhere [106].

GAUSSIAN UNITARIES AND SYMPLECTIC TRANSFORMATIONS   Unitary transformations are called Gaussian if and only if they map Gaussian states to Gaussian states. Such transformations have simple representations in phase space. They transform the operator $\vec{\chi}$ according to

$$\vec{\chi} \to S\vec{\chi} + \vec{d}, \tag{3.63}$$

where $S \in Sp(4, \mathbb{R})$ is a symplectic matrix

$$S\Omega S^T = \Omega \tag{3.64}$$

and $\vec{d}$ is a real vector. Note that symplectic transformations have unit determinant $\det S = \det S^{-1} = \det S^T = 1$ and transform the first two moments (3.59) and (3.60) according to

$$\vec{\chi} \to S\vec{\chi} + \vec{d}, \quad \gamma \to S\gamma S^T, \tag{3.65}$$

respectively. Examples include displacements generated by $D(\alpha)$, which only influence $\vec{d}$, orthogonal transformations $SS^T = \mathbb{1}$ (such as rotations) and squeezing $S = \text{diag}(a, 1/a)$ with $a > 0$. Starting from the vacuum state, any pure Gaussian state can be generated by applying one of these three transformations.

WILLIAMSON'S THEOREM   A central theorem in the context of Gaussian states is Williamson's theorem [107]. It states that for every covariance matrix γ there exists a symplectic transformation S diagonalizing it, i.e.

$$\gamma = S\Gamma S^T, \tag{3.66}$$

with $\Gamma = \text{diag}(\nu, \nu)$ and ν denoting the so-called symplectic eigenvalue. Note that ν is also the eigenvalue of Γ. Interestingly, a Gaussian state is pure if and only if there exists a symplectic S such that

$$\gamma = \frac{1}{2}SS^T, \tag{3.67}$$

showing that pure Gaussian states are characterized by the symplectic eigenvalue being $\nu = \frac{1}{2}$.

EXAMPLE: CANONICAL PHASE SPACE   We sketch the action of the canonical displacement operator $D(\alpha)$ on the vacuum state together with the uncertainty principle in Figure 3.3a. In Figure 3.3b, we compare the vacuum Husimi Q-distribution $\bar{Q}(x,k)$ with the marginal distributions $\bar{f}(x)$ and $\bar{g}(k)$, which are all of Gaussian form.

(a) Canonical displacement operator

(b) Canonical distributions of direct and joint measurements
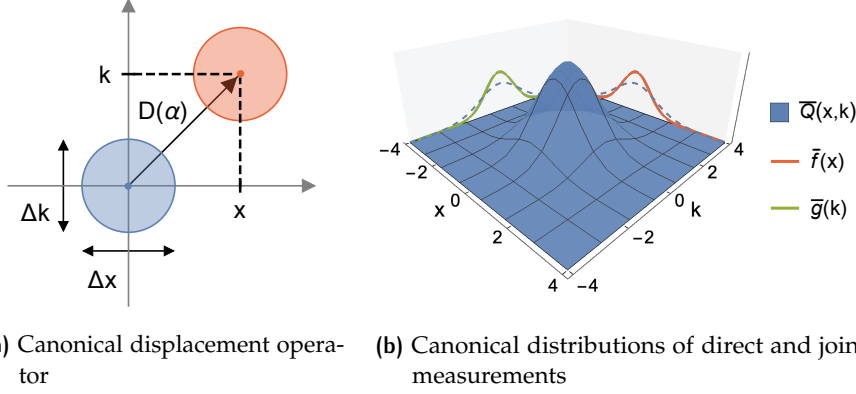
Figure 3.3: (a) Illustration of the canonical displacement operator and (b) the phase space distributions stemming from a joint or two direct measurements. As a result of the uncertainty principle, phase space distributions cannot be arbitrarily concentrated around one point. Also, one can read-off that the marginals of the Husimi Q-distribution (blue dashed lines) are broader than the probability densities $f(x)$ and $g(k)$. Note that the heights of the marginals differ from the height of the Husimi Q-distribution as a result of the factor $\frac{1}{2\pi}$ in the integral measure.

DISCRETE VARIABLES: DIRECT MEASUREMENTS   We close our analysis with the case of two *discrete* observables Y and Z, i.e. non-degenerate operators with finite and bounded spectra, typically corresponding to degrees of freedom of a finite-dimensional Hilbert space $D < \infty$. In contrast to the continuous variable case, we do not specify an algebra at this point, such that we can start from the general eigenvalue equations

$$Y \ket{y} = y \ket{y}, \quad Z \ket{z} = z \ket{z}, \tag{3.68}$$

where the eigenvectors now span an orthonormal basis, i.e.

$$\braket{y|y'} = \delta_{yy'}, \quad \braket{z|z'} = \delta_{zz'}. \tag{3.69}$$

These allow us to define two PVMs as $\{\ket{y}\bra{y}\}_y$ and $\{\ket{z}\bra{z}\}_z$, leading to measurement probabilities

$$p(y) = \braket{y|\rho|y}, \quad q(z) = \braket{z|\rho|z}. \tag{3.70}$$

Interestingly, discrete measurements allow for a well-defined notion of a post-measurement state without recording the result. For example, when measuring the observable Y, the state after measurement reads

$$\rho \to \rho_Y = \sum_{y \in \mathcal{Y}} p(y)\ket{y}\bra{y}. \tag{3.71}$$

According to (3.15), the von Neumann entropy of this state is equal to the Shannon entropy of the probability distribution $S(\rho_Y) = S(p)$. Similar considerations fail for a continuous variable X as

$$\rho \to \rho_X = \int_{x \in \mathcal{X}} \mathrm{d}x \, f(x) |x\rangle \langle x| \tag{3.72}$$

is *not* a trace-class operator and hence cannot represent any physical state. This is due to the fact that detecting a continuous variable would require infinite precision (and therefore also infinite energy), which is physically not feasible.

3.5 ENTROPIC UNCERTAINTY

SECOND MOMENT RELATIONS   The uncertainty principle lies at the heart of quantum theory and is maybe more famous than the theory itself. In simple words, it states that two non-commuting observables can *not* be measured or prepared simultaneously with arbitrary precision [108].

To quantify the uncertainty for measuring two arbitrary observables X and Y, various uncertainty relations have been derived. The most well-known and also rather general formulation of the uncertainty principle is due to Robertson [109]

"The more accurately the position is known, the less accurately is the momentum determined and vice versa." - Werner Heisenberg (translated) [108]

$$\sigma_x^2 \sigma_y^2 \geq \frac{1}{4} |\mathrm{Tr}\{\rho[X, Y]\}|^2, \tag{3.73}$$

where σ_x^2 denotes the variance of X, i.e.

$$\sigma_x^2 = \mathrm{Tr}\{\rho X^2\} - \mathrm{Tr}\{\rho X\}^2. \tag{3.74}$$

There exists also a stronger version of (3.73) with an extra term containing the anti-commutator of X and Y derived by Schrödinger [110] and Robertson [111]. We state it here for the special case of position X and momentum K, where it can be written as

$$\det \gamma \geq \frac{1}{4}. \tag{3.75}$$

Note that this relation is invariant under symplectic transformations as $\det S = \det S^T = 1$ and hence is tight for *all* pure Gaussian states.

For more than one mode, the uncertainty principle is typically expressed as the condition [112]

$$\gamma + \frac{i}{2}\Omega \geq 0, \tag{3.76}$$

or equivalently

$$\nu \geq \frac{1}{2}, \tag{3.77}$$

showing that γ is actually strictly positive-definite. For a single mode, the three relations (3.75), (3.76) and (3.77) are equivalent, while for more than one mode a multi-dimensional version of the former is implied by the two latter relations.

FLAWS OF VARIANCE Despite its simplicity, the formulation of the uncertainty principle in terms of variances or other second moments is not satisfactory from a quantum information theoretic perspective. First, the variance σ_x^2 is not a true measure of the uncertainty about the classical distributions $p(x)$ or $f(x)$ obtained after measuring the observable X several times. In fact, it is a second moment of this distribution, i.e.

$$\sigma_x^2 = \begin{cases} \sum_{x\in\mathcal{X}} p(x)x^2 - \left(\sum_{x\in\mathcal{X}} p(x)x\right)^2 & \text{for } x \text{ discrete,} \\ \int_{x\in\mathcal{X}} \mathrm{d}x\, f(x)x^2 - \left(\int_{x\in\mathcal{X}} \mathrm{d}x\, f(x)x\right)^2 & \text{for } x \text{ continuous,} \end{cases} \tag{3.78}$$

and hence lacks the information about all other moments of $p(x)$ or $f(x)$. Second, it shows counter-intuitive behavior in certain cases (see following example). Third, it can only be assigned to observables taking numerical values and fourth, it is not invariant under relabeling the measurement outcomes.

Hence, a more rigorous formulation in terms of entropy as *the* measure for uncertainty is appropriate, leading us to the notion of EURs. As classical information theory distinguishes between discrete and continuous variables, we treat those two cases separately in the following and start with the latter.

EXAMPLE: PARTICLE IN A BOX The counter-intuitive behavior of the variance is exemplified for a particle in a box in Figure 3.4, showing that the variance puts too much weight on the tails of a distribution [28, 113].

(a) Without walls: $\sigma_x \approx 0.29L$ (b) With walls: $\sigma_x \approx 0.5L$
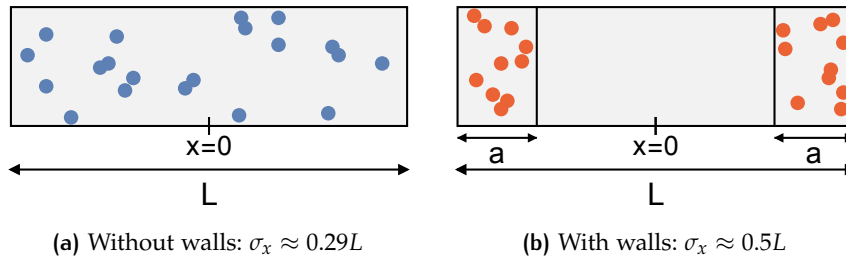
Figure 3.4: The position of a particle is measured inside a box of length L (each measurement is depicted by a single point). We distinguish the two cases (a) without walls and (b) with walls of width $a \ll L$. The variance of the position σ_x^2 is larger in the latter situation, although the uncertainty on the particle's position is lower on the right.

BIAŁYNICKI–BIRULA–MYCIELSKI EUR Following up on several works [114–116], the best-known EUR was derived for direct position and momentum measurements by BBM and reads [32]

$$S(f) + S(g) \geq 1 + \ln \pi. \tag{3.79}$$

Note that either of the two entropies can become negative. Interestingly, one can easily show that this relation implies the variance based relation for these variables

$$\sigma_x^2 \sigma_k^2 \geq \frac{1}{4}, \tag{3.80}$$

which is the well-known relation by Heisenberg and Kennard [108, 117] (see also [118]). Note that this relation follows from (3.75) by aligning the principal axes with the coordinate axes $\sigma_{xk} = 0$, showing that (3.75) is tighter than (3.80) if and only if $\sigma_{xk} \neq 0$.

Using that the probability density function maximizing the differential entropy for a given variance is a Gaussian leads to

$$\ln(2e\pi\sigma_x\sigma_k) = S(f_G) + S(g_G) \geq S(f) + S(g) \geq \ln e\pi. \tag{3.81}$$

Hence, the EUR in (3.79) is not only favorable from an information theoretic perspective, but is also *stronger* than the variance based relation (3.80).

In general, (3.79) is not invariant under rotations in phase space and is tight for all squeezed coherent states. It is an interesting open problem to formulate an EUR which is tight for all pure Gaussian states and reduces to (3.75) in the case of Gaussian states (see [29, 119, 120] and also [121, 122]).

In Chapter 9, we will derive the first EUR for quantum fields starting from (3.79).

FRANK–LIEB EUR Another EUR which can be applied to direct position and momentum measurements has been put forward by Frank and Lieb based on a work by Rumin [123], which reads [33]

$$S(f) + S(g) \geq \ln 2\pi + S(\rho). \tag{3.82}$$

In contrast to the state-independent bound of (3.79), the bound of the latter relation contains the von Neumann entropy $S(\rho)$, which accounts for the mixedness of the quantum state ρ. However, for pure states (3.82) is weaker than (3.79). In contrast to (3.79), the FL relation becomes tight in the infinite temperature limit.

WEHRL ENTROPY AND CANONICAL WEHRL–LIEB EUR The properties (3.52) and (3.53) allow us to associate an entropy with the Husimi Q-distribution, which is the *canonical Wehrl entropy* [22, 23]

$$S(Q) = -\int \frac{\mathrm{d}x\,\mathrm{d}k}{2\pi} Q(x,k) \ln Q(x,k). \tag{3.83}$$

This notion of an entropy in phase space fulfills an EUR known as the WL inequality [34, 124] (see also [125–128])

$$S(Q) \geq 1, \tag{3.84}$$

where equality is achieved if and only if the state under consideration is a pure coherent state projector $\rho = |\alpha\rangle\langle\alpha|$. Therefore, the canonical Wehrl entropy $S(Q)$ can be thought of as a *measure for localization* or *classicality* in phase space.

Note that (3.84) is, in contrast to (3.79), invariant under rotations, but not under squeezing. Note also that variants of WL inequality (3.84) have been proven for various algebras, including $SU(2)$ [124], symmetric $SU(K)$ [129] and $SU(1,1)$ [130], showing the generality of the phase space approach. Moreover, $S(Q)$ is an upper bound for the quantum entropy of the state $S(\rho)$, i.e.

$$S(\rho) \leq S(Q), \tag{3.85}$$

which renders it a coarse-grained entropy.

In Chapter 10, we investigate the Wehrl entropy further and compare the three EURs (3.79), (3.82) and (3.84) regarding their tightness.

STATE–DEPENDENT MAASSEN–UFFINK EUR   Based on the works [131, 132], the standard EUR for PVMs of discrete variables was formulated by Maassen and Uffink [30]. Here, we directly put forward the improved state-dependent version [31, 133]

$$S(p) + S(q) \geq -\ln c + S(\rho), \tag{3.86}$$

where

$$c = \max_{y,z}|\langle y|z\rangle|^2 \tag{3.87}$$

denotes the maximum overlap between any two basis vectors $|y\rangle$ and $|z\rangle$, and $S(\rho)$ improves the bound for mixed states. One may interpret c as the quantum incompatibility of the two measurement bases. In this sense, it plays a similar role as the commutator in Robertson's relation (3.73).

The MU EUR (3.86) can also be generalized for POVMs. It keeps its form when measuring the sets $\{E_y\}_y$ and $\{E_z\}_z$, with the quantum incompatibility c substituted by [12, 134–136]

$$c = \max_{y,z}\left\|\sqrt{E_y}\,E_z\,\sqrt{E_y}\right\|_\infty, \tag{3.88}$$

with $\|.\|_\infty$ being the infinity operator norm.

Note also that the relations (3.82) and (3.86) can be unified in a rather formal way as done by Frank and Lieb in [33] (see [137] for POVMs). This will be the starting point for a more direct unification in terms of relative entropies carried out in Chapter 8.

BEYOND THE SCOPE   Let us mention other directions in the wide field of EURs, which are not further discussed in this thesis.

First, refinements of the presented EURs were discussed extensively. For example, the MU EUR was strengthened in [135, 136], while refinements of the WL EUR are discussed in [138, 139].

Second, EURs have been formulated for bi- and tri-partite systems in the presence of (quantum) memory, where entropies are replaced by conditional entropies, see e.g. [31, 133, 140–145]. In this context, there are also reformulations in terms of mutual information, which are called information exclusion principles [136, 146, 147].

Third, besides the class of EURs presented so far, which are all concerned with preparation uncertainty in the sense that they quantify the incompatibility of measurements, there exists measurement EURs, which describe the disturbance of observables caused by subsequent measurements [148, 149].

Other topics include relations for energy and time [150–153], extensions to Rényi entropies [30, 154] and relations for complementary operator algebras [155–157].

3.6 ENTANGLEMENT

DEFINITION OF SEPARABILITY AND ENTANGLEMENT Quantum entanglement may be regarded as *the* feature distinguishing classical from quantum physics when at least two systems are considered. In the following, we examine a bipartition 12 of two systems 1 and 2 with Hilbert space $\mathcal{H}_{12} = \mathcal{H}_1 \otimes \mathcal{H}_2$.

We call the global state ρ_{12} a *product state* if it can be written as a tensor product of the local states, i.e.

$$\rho_{12} = \rho_1 \otimes \rho_2. \tag{3.89}$$

In this case, the two subsystems 1 and 2 are uncorrelated.

Furthermore, we say that ρ_{12} is *separable* if it can be written as a convex sum over product states, i.e.

$$\rho_{12} = \sum_i p(i) \left(\rho_1^i \otimes \rho_2^i \right), \tag{3.90}$$

where $p(i)$ is a probability distribution and $\{\rho_1^i\}_i$ and $\{\rho_2^i\}_i$ denote families of local states corresponding to subsystems 1 and 2, respectively. From an operational point of view, this means that the two subsystems 1 and 2 are classically correlated in the sense that the state ρ_{12} can be created by two parties with Local Operations and Classical Communication (LOCC) [159]. More precisely, if we imagine two parties Alice and Bob having access to only their subsystems 1 and 2, respectively, Bob may produce the probability distribution $p(i)$ and send it to Alice via classical communication. This is enough for both parties being able to prepare their local states, such that the global state will be of the form (3.90). Also, note that a globally pure

"The phenomenon of entanglement is the essential fact of quantum mechanics, the fact that makes it so different from classical physics. It brings into question our entire understanding about what is real in the physical world." - Leonard Susskind [158]

state is separable if and only if it is a product state of two pure local states in the sense of (3.89).

If a decomposition of the form (3.90) can *not* be found, ρ_{12} is said to be *entangled*. Then, the subsystems 1 and 2 exhibit quantum correlations, which discriminate quantum from classical physics. Operationally, this means that every entangled state can be used in order to execute some tasks with better performance than possible with any separable state [159] and hence becomes some kind of useful resource.

The mathematical theory of entanglement is mostly concerned with the separability problem: given some global state ρ_{12} or access to it via some measurement procedure, how can we *decide* whether this state is entangled or not? Often, it is also of great interest to ask: how can we *quantify* the amount of entanglement? We will briefly discuss common approaches in the following.

EXAMPLE: SET OF ENTANGLED STATES   The sets of separable and entangled states are sketched in Figure 3.5 with an emphasis put on the geometric meaning of purity.
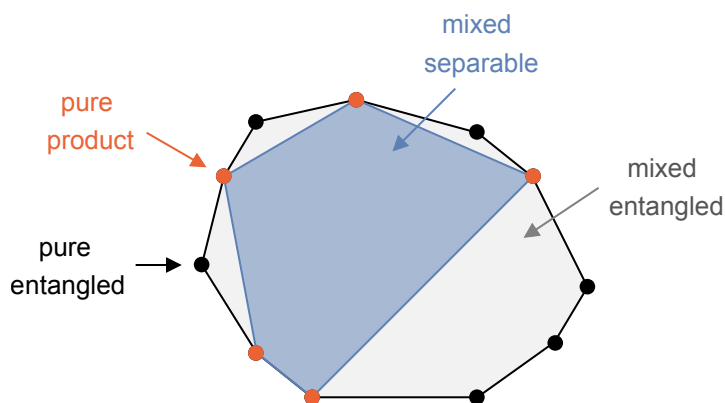
Figure 3.5: The set of mixed separable states (blue area) is the convex hull of the set of all pure product states (red points). As the set of all states is convex itself, pure product states as well as pure entangled states (black points) serve as extremal points. The leftover region (gray area) corresponds to the mixed entangled states.

ENTANGLEMENT MEASURES   Let us start with the question how entanglement can be quantified. To that end, we distinguish two cases, namely we consider the global state ρ_{12} to be either pure or mixed.

If the global state is pure $S(\rho_{12}) = 0$, the set of separable states is given by all product states of the form (3.89). In this case we can apply subadditivity of quantum entropies (cf. (3.23)) to find

$$\rho_{12} \text{ is } \begin{cases} \text{separable} & \Leftrightarrow & S(\rho_1) = S(\rho_2) = 0, \\ \text{entangled} & \Leftrightarrow & S(\rho_1) = S(\rho_2) > 0. \end{cases} \qquad (3.91)$$

In simple words, the latter condition means that entanglement is present if we know everything about the system as a whole, but not everything about its parts. Note also that the latter two statements can be reformulated with quantum conditional entropy becoming negative if and only if the state ρ_{12} is entangled as evident from the definition (3.19). Consequently, it is natural to measure the amount of entanglement by the quantum entropies $S(\rho_1) = S(\rho_2)$, which is why they are conveniently called *entanglement entropies*.

Instead, if the global state is mixed $S(\rho_{12}) > 0$, the situation becomes more involved. General entanglement measures are defined via a set of axioms, which follow the operational approach to entanglement theory. Although there are numerous different characterizations, we follow one of the most common approaches based on the following four axioms [69, 159]. We say that $E(\rho_{12})$ is an *entanglement measure* for subsystems 1 and 2 if it

1. outputs a real non-negative number

$$E(\rho_{12}) \geq 0 \text{ for all } \rho_{12} \in \mathcal{B}(\mathcal{H}_{12}), \tag{3.92}$$

2. becomes zero if and only if the state is separable

$$E(\rho_{12}) = 0 \Leftrightarrow \rho_{12} \text{ separable}, \tag{3.93}$$

3. does not increase under any LOCC represented by $\mathcal{N}$

$$E(\rho_{12}) \geq E(\mathcal{N}(\rho_{12})), \tag{3.94}$$

4. reduces to the entanglement entropy for pure states

$$E(\rho_{12}) = S(\rho_1) = S(\rho_2) \text{ for all } \rho_{12} = |\psi\rangle\langle\psi|. \tag{3.95}$$

One may regard the third axiom as central to this characterization and it should be noted that this requirement can be motivated from the fact that LOCCs can cannot create entanglement. Hence, a reasonable entanglement measure should not increase when LOCCs are applied to the global state. Furthermore, the third axiom immediately implies that entanglement measures have to be invariant under local unitary transformations. Note also that if the fourth axiom is not met, one we speak of an *entanglement monotone*.

EXAMPLE: RELATIVE ENTROPY OF ENTANGLEMENT   Several quantities which meet these four axioms exist. Exemplary, we discuss the Relative Entropy of Entanglement (REoE), which is defined as the optimization problem [69, 160]

$$E_{\mathrm{REoE}}(\rho_{12}) = \inf_{\tilde{\rho}_{12} \text{ sep.}} S(\rho_{12} \| \tilde{\rho}_{12}), \tag{3.96}$$

where the minimization is over all separable model states $\tilde{\rho}_{12}$.

From the definition of the quantum relative entropy it follows immediately that the REoE fulfills axioms one and two, while its monotonicity property (3.18) ensures the third axiom to be valid. Moreover, for globally pure states it was proven that the REoE reduces to the entanglement entropy in [160].

Note the close resemblance to the quantum mutual information, which follows for $\tilde{\rho}_{12} = \rho_1 \otimes \rho_2$, such that the REoE may be regarded as the mixed state generalization of the quantum mutual information. By definition they are related via $E_{\mathrm{REoE}}(\rho_{12}) \leq I(\rho_1 : \rho_2)$.

At last, let us mention that the REoE (and all other entanglement measures) are often NP-hard to compute and therefore as hard as the separability problem itself [161, 162].

ENTANGLEMENT DETECTION VIA SEPARABILITY CRITERIA As a result from our discussion of entanglement measures, it is evident that the pure state case is trivial as the entanglement entropy provides a unique measure and allows to decide easily whether the global state is entangled or not. In contrast, mixed state entanglement turned out to be NP-hard to quantify. Hence, we may formulate criteria which are simple to evaluate on the cost of returning inconclusive results for some states, allowing us to characterize at least some entangled states.

In the following, we speak of a separability criterion as a condition, often formulated for measurable quantities, which is *met by all separable states*. Therefore, the violation of a separability criterion by some state certifies that this state is entangled. On the other hand, if some state fulfills the criterion, we can not say whether this state is entangled or not. In this sense, a separability criterion is a necessary condition for a state to be separable, but is typically not sufficient. Hence, the violation of the criterion is a sufficient condition for entanglement. However, in some rare cases, certain separability criteria turn out to be sufficient conditions for separability as well. In these cases, the separability problem is solved.

From an information theoretic perspective, there is a very simple criterion based on the quantum conditional entropy, which states that

$$S(\rho_1|\rho_2), S(\rho_2|\rho_1) \geq 0. \tag{3.97}$$

for all separable states. It follows from generalizing the trivial argument for pure states to mixed states via concavity. Hence, if a quantum state has negative quantum conditional entropy, it is entangled [35]. Note again that this condition is also sufficient for pure states.

PPT CRITERION We proceed with the probably most well-known separability criterion, which is the *Peres-Horodecki criterion*, also known as Positive Partial Transpose (PPT) criterion. It states that any separable state ρ_{12} has PPT [163, 164]

$$\rho_{12} \xrightarrow{T_2} (\mathbb{1} \otimes T_2) \rho_{12} \geq 0, \tag{3.98}$$

"It would not make sense for a mechanic to say, 'I know everything about your car but unfortunately I can't tell you anything about any of its parts.' But that's exactly what Einstein explained to Bohr — in quantum mechanics, one can know everything about a system and nothing about its individual parts — [...]." - Leonard Susskind [158]

i.e. the operator $(\mathbb{1} \otimes T_2)\,\rho_{12}$ resulting from using the partial transpose on subsystem 2 (or analogously on subsystem 1), is non-negative. The statement is a simple consequence from the definition (3.90). For separable states (3.90), applying a partial transposition results in

$$\rho_{12} \xrightarrow{T_2} \sum_i p(i)\left(\rho_1^i \otimes \left(\rho_2^i\right)^{T_2}\right). \tag{3.99}$$

Since $\left(\rho_2^i\right)^{T_2}$ has the same eigenvalues as $\left(\rho_2^i\right)$, the resulting operator needs to be non-negative for separable states.

Interestingly, the PPT criterion is also *sufficient* for finite Hilbert spaces with dimensions 2×2 and 2×3 [164]. In higher dimensions, the PPT criterion is only necessary. In this case, there exist entangled states with positive partial transpose, which are called bound entangled states [165].

THE LINK BETWEEN ENTANGLEMENT AND UNCERTAINTY   The partial transpose does not change the normalization and hermicity of the initial state ρ_{12} and hence $\mathbb{1} \otimes T_2$ is a PTP map. If the state under consideration is separable, it becomes a CPTP map and the resulting operator is a density operator representing a physical quantum state as defined in Section 3.1 (see also [42] for the deep connection between separability criteria and PTP maps). As such, all derived distributions of non-commuting observables are constrained by the uncertainty principle [166]. As a result, *any* uncertainty relation, when applied to a partially transposed operator, can be used to infer a separability criterion.

However, not all choices of observables are practical. If an uncertainty relation exhibits symmetries, the resulting criteria are sometimes fulfilled by all entangled states, too. Therefore, one has to come up with appropriate observables, such that the derived criteria are violated by at least a couple of entangled states.

PPT CRITERION FOR CONTINUOUS VARIABLES   We will discuss a few of such criteria for continuous variable systems in the following. Interestingly, the partial transpose has a simple geometric interpretation in phase space: it corresponds to a mirror reflection in one local phase space, i.e. for every quasi-probability distribution in phase space the partial transpose transforms the coordinates as (see also Section 13.1)

$$\vec{\chi}_{12} = (x_1, k_1, x_2, k_2) \xrightarrow{T_2} \Lambda\vec{\chi}_{12} = (x_1, k_1, x_2, -k_2), \tag{3.100}$$

with the matrix

$$\Lambda = \mathrm{diag}(1, 1, 1, -1). \tag{3.101}$$

Several separability criteria can be obtained, depending on which observables and uncertainty relations are considered.

SECOND MOMENT CRITERIA FOR CONTINUOUS VARIABLES When employing second moment uncertainty relations, one can end up with one of the following three well-known criteria. Simon [44] started with (3.76) constraining the global phase space covariance matrix γ_{12} and obtained the separability criterion

$$\gamma'_{12} + \frac{i}{2}\Omega \geq 0, \tag{3.102}$$

where

$$\gamma'_{12} = \Lambda\gamma_{12}\Lambda \tag{3.103}$$

denotes the global covariance matrix corresponding to the partially transposed operator $(\mathbb{1} \otimes T_2)\rho_{12}$.

Another approach is based on considering non-local observables in the spirit of Einstein-Podolsky-Rosen (EPR) [167]

$$X_\pm = |a|X_1 \pm \frac{1}{a}X_2, \quad K_\pm = |a|K_1 \pm \frac{1}{a}K_2, \tag{3.104}$$

with $a \neq 0$ being a real number (which accounts for local squeezing). These operators fulfill the commutation relations

$$[X_\pm, K_\pm] = a^2 + \frac{1}{a^2}, \quad [X_\pm, K_\mp] = a^2 - \frac{1}{a^2}. \tag{3.105}$$

The variances $\sigma^2_{x_\pm}$ and $\sigma^2_{k_\pm}$ of the measured distributions $f_\pm \equiv f(x_\pm)$ and $g_\pm \equiv g(k_\pm)$, respectively, are constrained by the uncertainty principle in the form of Robertson's relation (3.73). Now, applying the PPT criterion according to (3.100) results in $k_\pm \to k_\mp$. Then, we end up with the Mancini-Giovannetti-Vitali-Tombesi (MGVT) separability criteria [45, 46]

$$\sigma_{x_\pm}\sigma_{k_\mp} \geq \frac{1}{2}\left(a^2 + \frac{1}{a^2}\right) \geq 1. \tag{3.106}$$

A weaker set of criteria, the Duan-Giedke-Cirac-Zoller (DGCZ) criteria, can be obtained by bounding the left hand side from above, resulting in [43]

$$\sigma^2_{x_\pm} + \sigma^2_{k_\mp} \geq a^2 + \frac{1}{a^2} \geq 2, \tag{3.107}$$

where we used that $bc \leq \frac{1}{2}(b^2 + c^2)$ as a consequence of $(b-c)^2 \geq 0$. In both cases we also gave the minimum of the corresponding bound, which leads to weaker criteria in general.

Note that the latter two criteria were originally derived without the use of the PPT criterion. Note also that the second moment criteria by Simon (3.102) and DGCZ (3.106) are formally equivalent and that all three criteria coincide for Gaussian states, in which case they are also *sufficient* criteria for separability.

ENTROPIC CRITERIA FOR CONTINUOUS VARIABLES   What the second moment criteria have in common is that they are all easy to evaluate, but often return inconclusive statements for entangled but non-Gaussian states. To overcome this problem, criteria involving higher moments have been formulated [168, 169].

In contrast, the modern quantum information theoretic approach relies on entropic criteria, which are generically more sensitive in the non-Gaussian regime. For example, when considering the non-local operators (3.104) with $a = 1$ and using the BBM EUR (3.79) instead of a second moment uncertainty relation, one ends up with the entropic criteria by Walborn-Taketani-Salles-Toscano-de Matos Filho (WTSTD) [47] (see also [48, 49])

$$S(f_\pm) + S(g_\mp) \geq 1 + \ln \pi. \tag{3.108}$$

Again, the original derivation in [47] does not utilize the PPT criterion. Extensions to Rényi entropies can be found in [49], while differential conditional entropies are discussed in [51, 144, 145].

In Chapter 13 we will use the PPT criterion in phase space to derive a general family of beyond-entropic separability criteria.

BEYOND THE SCOPE   Let us close with a short list of topics connected to entanglement theory, which will not be discussed any further in this thesis.

First, there are many more separability criteria, especially in the context of discrete variable quantum systems. A notable mention is the reduction criterion [170, 171]. See also [142, 143, 172, 173] for various entropic criteria for discrete variables.

Second, there is the large field of multipartite entanglement, i.e. the study of entanglement between more than two parts of a system [42, 65]. This includes other interesting peculiarities of the phenomenon entanglement, for example the monogamy of entanglement.

Third, entanglement is the defining feature in quantum metrology, which allows one to perform experiments with higher precision than any classical analog [174].

Fourth, entanglement is key for quantum computing as it is responsible for the performance gain over classical computations when applied to suitable tasks [14, 15].

4 | QUANTUM FIELDS

We proceed with QFT, which underpins the modern construct of ideas in physics. Especially in recent years, information theoretic methods have been applied extensively to field-theoretic problems, for example in the context of black holes [40, 175–178], the famous information paradox [16–18], holography [41, 179–181], condensed matter physics [182–184] or high energy physics [185–191].

In the following, we briefly introduce QFT for a scalar field within the functional integral formalism [192] and discuss the role of entanglement in QFT based on [182, 183, 193–196].

4.1 FUNCTIONAL INTEGRAL FORMULATION

CLASSICAL FIELDS In contrast to classical mechanics, classical fields describe *infinitely* many degrees of freedom. More precisely, a real scalar field $\phi(x)$ attains some real value at every spacetime point x. Heuristically, one may think of such a field as the limit of infinitely many oscillator modes. This can be made explicit by working in terms of a lattice theory, such that we deal with a field $\phi(x_{i_1 \ldots i_d})$ evaluated at discretized spacetime points $x_{i_1 \ldots i_d} = \epsilon \, (i_1, \ldots, i_d)^T$, where ϵ is the lattice spacing between two neighboring modes and i_j is integer-valued for all $1 \leq j \leq d$. Then, we get back the continuous field in the *continuum limit*, i.e. $x_{i_1 \ldots i_d} \to x$ and $\phi(x_{i_1 \ldots i_d}) \to \phi(x)$ for $\epsilon \to 0$ and $N \to \infty$, where N is the number of modes per site.

A field theory is further specified by an action

$$S = \int \mathrm{d}t \, L = \int \mathrm{d}^d x \, \mathcal{L}, \tag{4.1}$$

where $L = L(\phi, \partial_t \phi)$ denotes the Lagrangian and $\mathcal{L}$ the corresponding density. Typically, $\mathcal{L}$ is a local function of the underlying field $\phi(x)$ and derivatives thereof in the sense that causality has to be respected, which is especially the case if the interactions are local. On the lattice, this requirement translates into nearest-neighbor interactions.

Next, we introduce the conjugate momentum field $\pi(x)$ via

$$\pi(x) = \frac{\partial \mathcal{L}}{\partial(\partial_t \phi)}, \tag{4.2}$$

which allows us to define the Hamiltonian density $\mathcal{H} = \mathcal{H}(\phi, \pi)$ as the Legendre transform of the Lagrangian density $\mathcal{L}$, i.e.

$$\mathcal{H} = \pi(x)\partial_t \phi(x) - \mathcal{L}. \tag{4.3}$$

Note that one obtains the Hamiltonian H by integrating the latter over space

$$H = \int \mathrm{d}^{d-1}x \, \mathcal{H}. \tag{4.4}$$

Although it is sometimes convenient to work within the Hamiltonian formalism, it is important to note that for relativistic theories the Lagrangian density $\mathcal{L}$ is Lorentz-invariant, i.e. transforms as a scalar under Lorentz transformations, while the Hamiltonian density $\mathcal{H}$ is not. This can be seen from (3.9), which shows that the Hamiltonian H generates translations in time and hence singles out a time direction. Therefore, despite the fact that both formulations are covariant (in the sense of form-invariant), the symmetries become *manifest* only when working with $\mathcal{L}$.

EXAMPLE: CONTINUUM LIMIT   We sketch the continuum limit $\epsilon \to 0, N \to \infty$ of a $d = 1 + 0$-dimensional lattice theory in Figure 4.1.

(a) $\epsilon = 0.5, N = 10$ **(b)** $\epsilon = 0.25, N = 20$ **(c)** $\epsilon \to 0, N \to \infty$
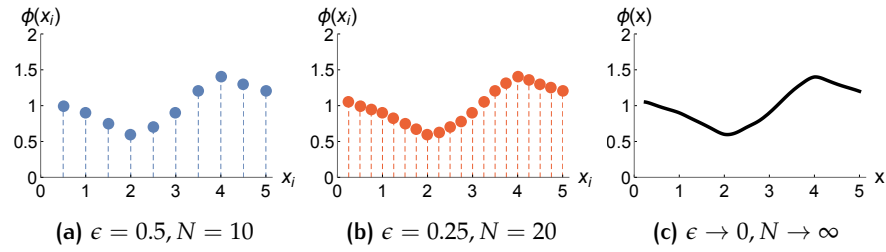
Figure 4.1: Continuum limit $\epsilon \to 0, N \to \infty$ for a scalar field $\phi(x)$ on a lattice with (in-)decreasing lattice spacing ϵ (number of modes N) from (a) to (c).

FROM CLASSICAL TO QUANTUM FIELDS I: CANONICAL QUANTIZATION
The transition from classical to quantum fields can be made in two ways. In the canonical formalism, the fields are promoted to field *operators* $\phi(x) \to \boldsymbol{\phi}(x), \pi(x) \to \boldsymbol{\pi}(x)$ and one imposes commutation relations at every two points in spacetime

$$[\boldsymbol{\phi}(x), \boldsymbol{\pi}(x')] = \delta^{(d)}(x - x'). \tag{4.5}$$

The field operators can be expanded in terms of creation and annihilation operators, $\boldsymbol{a}_{\vec{k}}^{\dagger}$ and $\boldsymbol{a}_{\vec{k}}$, respectively, where $\vec{k}$ denotes a momentum. For example, we obtain in the Schrödinger picture at a constant time t

$$\boldsymbol{\phi}(\vec{x}) = \int \frac{\mathrm{d}^{d-1}k}{(2\pi)^{d-1}} \frac{1}{\sqrt{2\omega(k)}} \left(\boldsymbol{a}_{\vec{k}} \, e^{i\vec{k}\vec{x}} + \boldsymbol{a}_{\vec{k}}^{\dagger} \, e^{-i\vec{k}\vec{x}} \right), \tag{4.6}$$

where $\omega(k)$ denotes the dispersion relation, which is specified by the concrete theory of interest. This decomposition allows for the useful interpretation that an excitation of the quantum field $\boldsymbol{\phi}(x)$ corresponds to a free particle with definite momentum $\vec{k}$. To calculate typical observables such as correlation functions, one may prefer the following second option.

FROM CLASSICAL TO QUANTUM FIELDS II: FUNCTIONAL INTEGRALS

In the functional integral formalism, one instead takes Feynman's path integral approach in QM to the field-theoretic level, i.e. one constructs a *functional integral* over all field configurations with a complex-valued measure. We introduce the corresponding partition function as a functional integral over all possible field configurations (constrained by the state of interest)

$$Z = \int \mathcal{D}\phi \, e^{iS} = \lim_{\epsilon \to 0} \left(\prod_{i_1,\ldots,i_d \in \mathbb{Z}} \int \mathrm{d}\phi(x_{i_1\ldots i_d}) \, e^{iS} \right), \tag{4.7}$$

where the functional integral $\int \mathcal{D}\phi$ is defined via the continuum limit $\epsilon \to 0$ and the complex-valued weighting factor is given by e^{iS}.

In this formalism, correlation functions follow directly from the partition function Z. If one introduces a source term $J(x)$, the latter becomes

$$Z[J] = \int \mathcal{D}\phi \, e^{iS + \int \mathrm{d}^d x J(x)\phi(x)}. \tag{4.8}$$

From this expression, correlation functions can be computed via *functional derivatives* at vanishing sources. For example, the field expectation value can be computed from

$$\langle \phi(x) \rangle = \frac{1}{Z[J]} \frac{\delta Z[J]}{\delta J(x)} \Big|_{J=0} = \frac{\int \mathcal{D}\phi \, \phi(x) \, e^{iS}}{\int \mathcal{D}\phi \, e^{iS}}, \tag{4.9}$$

with the functional derivative being defined through

$$\frac{\delta \phi(x)}{\delta \phi(x')} = \delta^d(x - x'). \tag{4.10}$$

Often, it is convenient to perform an analytic continuation (also called Wick rotation), which transforms the complex-valued weighting factors e^{iS} into real-valued factors e^{-S}, allowing for convergence of the functional integral. The analytic continuation is heuristically motivated by the fact that the Euclidean and Minkowskian metrics are equivalent if the time coordinate τ of the former is imaginary and related to the latter via $t \to -i\tau$, such that $S \to -iS$. Consequently, one refers to the latter two as Minkowskian (real time) and Euclidean (imaginary time) formalism, respectively.

The connections between the Minkowskian and Euclidean theories are studied in detail in the context of the Osterwalder-Schrader theorem [197–199]. Interestingly, they are very hard to establish formally. For example, it is one of the millennium problems to prove the existence of such a connection for gauge theories in $d = 4$ spacetime dimensions in full mathematical rigor [200].

CAUCHY HYPERSURFACES AND QUANTUM STATES   In QM, we considered quantum states at every instance of time $\rho = \rho(t)$ (see (3.7)). However, if we consider a quantum field on some generic background geometry specified by a metric $g_{\mu\nu}$, we have to work with a foliation of spacetime instead. This corresponds to a family of so-called *Cauchy hypersurfaces* $\Sigma(t)$, which are $(d-1)$-dimensional submanifolds of spacetime equipped with a time-like normal vector $n_\mu(x)$ at every point $x \in \Sigma$. The parameter t plays the role of a time coordinate and sometimes indeed corresponds to a time measured by some particular observer. In this sense, time evolution can be generalized to an evolution from an initial Cauchy hypersurface to a final hypersurface, i.e. $\Sigma(t_i) \to \Sigma(t_f)$.

Then, a quantum state can be associated with every hypersurface $\Sigma(t)$, i.e. $\rho(t)$. Often, it is convenient to express the state (or other operators) in terms of its matrix representation $\rho[\phi_+, \phi_-]$ with respect to the fields $\phi_+(x)$ and $\phi_-(x)$ evaluated on the hypersurface Σ, i.e. $x \in \Sigma$. As we will see in Chapter 9, the diagonal elements of this matrix $\rho[\phi, \phi]$ are of special interest to entropic uncertainty in QFT.

TIME EVOLUTION OF A CLOSED SYSTEM   Let us consider a closed system and a unitary time evolution from a state $\rho(t_i)$ on one hypersurface Σ_i to the next, i.e. to $\rho(t_f)$ on Σ_f. In the aforementioned matrix representation, we can write the unitary time evolution equation (3.8) as

$$\rho[\phi_{f+}, \phi_{f-}](t_f) = \int \mathcal{D}\phi_{i+} \int \mathcal{D}\phi_{i-} \, U[\phi_{f+}, \phi_{i+}](t_i, t_f) \tag{4.11}$$
$$\rho[\phi_{i+}, \phi_{i-}](t_i) U^\dagger[\phi_{i-}, \phi_{f-}](t_f, t_i),$$

wherein the unitary time evolution operator from Σ_i to Σ_f is given by the functional integral

$$U[\phi_f, \phi_i](t_i, t_f) = \int \mathcal{D}\phi \, e^{iS[\phi]} \tag{4.12}$$

with the boundary conditions $\phi(x) = \phi_i(x)$ on Σ_i and $\phi(x) = \phi_f(x)$ on Σ_f understood. The functional integral (4.11) together with (4.12) is referred to as the Schwinger-Keldysh double time path, as the two unitary evolution operators in (4.11) correspond to an evolution forward (with U) and backwards (with $U^\dagger$) in time [201, 202].

Starting from that, we will discuss information theoretic constraints on the time evolution of an open quantum system in Chapter 7.

EXAMPLE: CAUCHY HYPERSURFACES   We sketch flat and curved Cauchy hypersurfaces for a $d = 1+1$-dimensional theory in Figure 4.2a and Figure 4.2b, respectively.

(a) Flat hypersurfaces   **(b)** Curved hypersurfaces
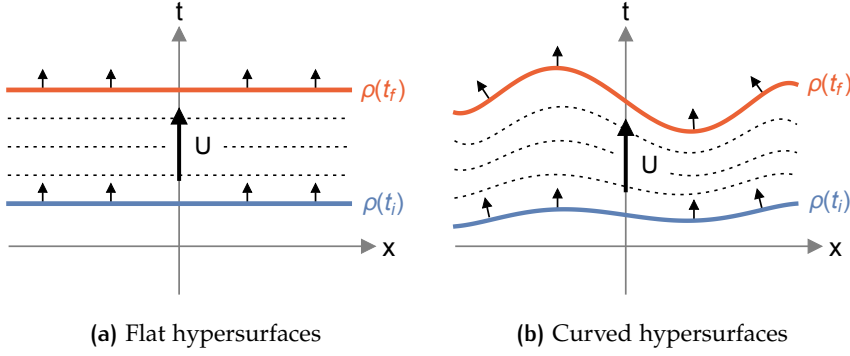
Figure 4.2: Families of Cauchy hypersurfaces with time-like normal vectors (small black arrows) representing time evolution from a quantum state $\rho(t_i)$ (blue line) to $\rho(t_f)$ (red line) via the unitary operator U (large black arrow). In (a), we consider the special case of flat hypersurfaces, i.e. $g_{\mu\nu} = \eta_{\mu\nu}$, such that the normal vector equals the unit vector in time direction at every point. In (b), we picture a more general situation in the presence of curvature.

FIELDS AT EQUILIBRIUM For later purposes, we also discuss quantum fields at finite temperature. In a field-theoretic setting, the thermal density operator (3.28) can still be written as

$$\rho_c = \frac{1}{Z_c} e^{-\beta H}, \tag{4.13}$$

where the Hamiltonian operator H is now a local integral over the Hamiltonian density operator $\mathcal{H}$ in analogy to (4.4). Note that more generally, $\mathcal{H}$ is the time-time component of the energy-momentum tensor $T_{\mu\nu}$. The latter will appear in the covariant generalization of (4.13) to arbitrary Cauchy hypersurfaces, which will be the starting point for the discussion of relativistic fluid dynamics in Chapter 7.

Moreover, let us briefly discuss the thermal state in the functional integral formalism. Comparing the thermal density operator to the unitary time evolution operator (3.9) reveals that the two are equivalent if we set $t_f - t_i \equiv 0 - i\beta$. Hence, from (4.12) we obtain the functional integral representation of the thermal state

$$\rho_c[\phi_+, \phi_-] = \frac{1}{Z_c} \int \mathcal{D}\phi \, e^{iS[\phi]} \tag{4.14}$$

with an action

$$S[\phi] = \int_0^{-i\beta} dt \int d^{d-1}x \, \mathcal{L}[\phi] \tag{4.15}$$

and boundary conditions $\phi(x) = \phi_+(x)$ for $t = 0$ and $\phi(x) = \phi_-(x)$ for $t = -i\beta$, which imply that the functional integral goes over a torus geometry in time direction t. Note that we get back the vacuum result for $\beta \to \infty$.

4.2 ENTANGLEMENT FOR QUANTUM FIELDS

BIPARTITION IN A QFT In quantum theory, entanglement is studied for systems with finitely many degrees of freedom. For example, one may consider two spins or two harmonic oscillator modes. When working with quantum fields, one has to deal with infinitely many degrees of freedom, which requires a more careful treatment of the phenomenon entanglement, as it comes with non-renormalizable UV divergences for *several* quantities of interest.

To access these peculiarities, let us consider again a regularized quantum field on some space-like hypersurface Σ. In this situation, the inverse lattice spacing $1/\epsilon$ provides an UV-regulator. In this sense, the continuum limit $\epsilon \to 0$ takes into account arbitrarily high energy scales, while a finite ϵ cuts off energy scales above $1/\epsilon$. Also, in the regularized theory, every discrete spatial position $x_{i_2\ldots i_d} \in \Sigma$ is endowed with a single mode and a corresponding Hilbert space $\mathcal{H}_{x_{i_2\ldots i_d}}$. Depending on the theory of interest these Hilbert spaces are finite (for examples for spins) or infinite (for oscillator modes).

Now, the hypersurface Σ may be divided into two regions in several ways. In this work, we only consider a finite spatial region 1 and its complement 2 such that $\Sigma = 1 \cup 2 \equiv 12$ and $1 \cap 2 = \emptyset$. Alternatively, one could study two half-infinite regions 1 and 2, for example two half-lines in $d = 1+1$ spacetime dimensions. For the corresponding Hilbert spaces we have a decomposition of the form

$$\mathcal{H}_{12} = \mathcal{H}_1 \otimes \mathcal{H}_2 \tag{4.16}$$

with

$$\mathcal{H}_1 = \bigotimes_{x_{i_2\ldots i_d} \in 1} \mathcal{H}_{x_{i_2\ldots i_d}}, \quad \mathcal{H}_2 = \bigotimes_{x_{i_2\ldots i_d} \in 2} \mathcal{H}_{x_{i_2\ldots i_d}}. \tag{4.17}$$

Interestingly, the continuum limit $\epsilon \to 0$ leads to severe problems already on the level of the Hilbert space decomposition (4.16). In fact, such a decomposition does *not* exist in a strict sense for $\epsilon \to 0$.

Mathematically, this is due to the Reeh-Schlieder theorem [203]. In simple words, it states that operators localized to any region still suffice to generate all other operators, indicating that there is entanglement present between observables in any two regions.

Physically, the problems come from the presence of a (sharp) surface between the two regions 1 and 2, which is the so-called *entangling surface*. In our case, this corresponds to the boundary $\partial 1$ of region 1, which is typically assumed to be smooth. In the limit $\epsilon \to 0$, arbitrarily high energy modes are present, which are localized to an arbitrarily small region around the entangling surface $\partial 1$. Hence, a bipartite splitting of the Hilbert space becomes ambiguous.

To be on the safe side, we keep thinking in terms of a regularized theory in the following, such that the global Hilbert space can indeed

be decomposed according to (4.16). Starting from a global state ρ_{12}, we associate local states in analogy to (3.6)

$$\rho_1 = \mathrm{Tr}_2\{\rho_{12}\}, \quad \rho_2 = \mathrm{Tr}_1\{\rho_{12}\}, \tag{4.18}$$

with the regions 1 and 2, respectively. Based on that, we may ask how much the two regions 1 and 2 are entangled by computing the corresponding entanglement entropies.

EXAMPLE: ENTANGLEMENT IN QM VS. QFT   We illustrate bipartitions for two (a) quantum spins in Figure 4.3a and (b) spatial regions in a lattice field theory Figure 4.3b.

(a) Two entangled spins   **(b)** Two entangled regions
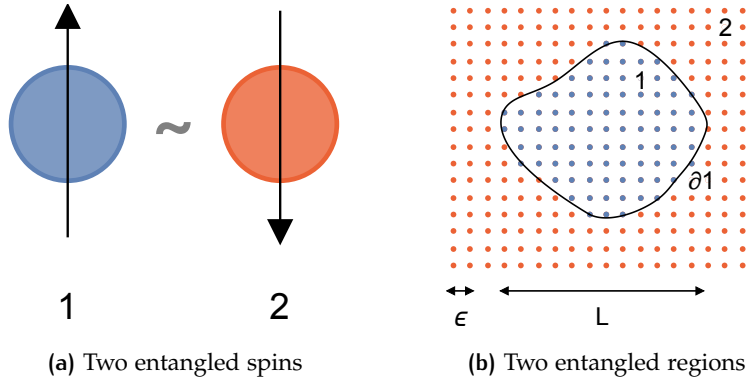
Figure 4.3: In quantum theory, entanglement is present between well-separated objects, e.g. between two quantum spins 1 and 2 sketched in (a). For quantum fields regularized on a lattice with lattice constant ϵ in (b), we consider a region 1 of typical size L and its complement 2, which are strongly entangled across the boundary $\partial 1$.

VOLUME LAW FOR TYPICAL STATES   To appreciate the significance of the area law discussed later, let us start with a naive estimation of the entanglement entropy for typical pure states, in which case the entanglement entropy provides a meaningful measure for entanglement.

We consider the volume of the enclosed region 1 to be of the order $V \sim L^{d-1}$ (cf. Figure 4.3b), i.e. we approximate 1 by a ball-shaped region. Second, we assume the size of 1 to be much larger than the cutoff-scale, i.e. $L \gg \epsilon$, but still small compared to the size of the full system 12. Often quantum fields extend up to infinity, such that the latter assumption reads $L \ll \infty$.

In such a situation we have that the Hilbert space $\mathcal{H}_1$ associated with 1 is much smaller than $\mathcal{H}_2$, allowing us to apply Page's theorem. It states that for a random global pure state ρ_{12} constructed from the unitarily invariant Haar measure, the local state ρ_1 is close to being maximally entangled [204–206]. Hence, we end up with

$$S(\rho_1) \sim \log D_1 \sim (L/\epsilon)^{d-1}, \tag{4.19}$$

which is the so-called *volume law* of entanglement entropy [183, 196]. We observe that the entanglement entropy of a typical state is proportional to the number of modes enclosed in 1 and hence becomes UV-divergent in the continuum limit $\epsilon \to 0$.

AREA LAW FOR PHYSICAL STATES However, in QFTs one is typically *not* concerned with random states, but rather with physical states, such as the vacuum, low-lying excitations or thermal states. Moreover, constraints come in when considering local QFTs, which is generically the case for relativistic theories, but also for non-relativistic theories with nearest-neighbor interactions in the lattice picture. In these cases, one may heuristically argue that entanglement between 1 and 2 is mainly caused by the constituents close to the entangling surface $\partial 1$. As a consequence, the entanglement entropy should in leading order be proportional to the area of the boundary region $\partial 1$.

The modes close to the boundary live in the UV-regime and hence the dominant contributions to the entanglement entropy come from UV-modes on small length scales. On such scales, the monotonicity of quantum relative entropy (3.18) implies that the vacuum is indistinguishable from any other finite energy state. Therefore, the entanglement entropy needs to have a rather universal form.

To motivate the following result for the entanglement entropy (we loosely follow [41, 207]), let us consider a theory, where no additional scales besides the size L of region 1 and the lattice constant ϵ are involved (cf. Figure 4.3b). We make the following ansatz for the entanglement entropy

$$S(\rho_1) \sim \int_{\partial 1} \mathrm{d}^{d-2}x \sqrt{h}\, F[K, h], \tag{4.20}$$

wherein x are coordinates, h the induced metric and K the extrinsic curvature on the entangling surface $\partial 1$, while F denotes some functional. We may expand the functional F in terms of the extrinsic curvature K. To that end, we note that $K \sim \frac{1}{L}$, which together with the fact that the entropy is a dimensionless quantity implies

$$S(\rho_1) \sim \frac{L^{d-2}}{\epsilon^{d-2}} + ... + \frac{L}{\epsilon} + \ln\frac{L}{\epsilon} + \text{const.} \tag{4.21}$$

More generally, one finds that the entanglement entropy of a physical state fulfills a so-called *area law* [183, 194, 195]

$$S(\rho_1) = \frac{g_{d-2}(\partial 1)}{\epsilon^{d-2}} + ... + \frac{g_1(\partial 1)}{\epsilon} + g_0(\partial 1)\ln\epsilon + \tilde{S}, \tag{4.22}$$

wherein $g_i(\partial 1)$ are local coefficients on the boundary $\partial 1$ depending on the theory, the state and the geometric form of the region 1 under consideration, while $\tilde{S}$ is a finite part often depending on the volume V. We observe that the leading order UV-divergence is proportional to the area and not to the volume as estimated in (4.19). This shows that

physical states occupy a very special region in the global Hilbert space $\mathcal{H}_{12}$. Also, one should note that the lowest order coefficient $g_0(\partial 1)$ is universal, i.e. independent of the regularization scheme, which is not the case for all other coefficients, while $\tilde{S}$ can contain universal as well as non-universal expressions.

Let us emphasize that although the area law (4.22) holds for all states, it only allows for a clear statement about entanglement if the global state ρ_{12} is pure. However, in this case, we need to have $S(\rho_1) = S(\rho_2)$, such that we can constrain (4.22) further extending on our heuristic argument from before. As the complement 2 comes with a sign change in the extrinsic curvature, only even powers in $\frac{L}{\epsilon}$ are allowed, and hence the area law does only contain even powers too. Also, note that the term $g_0(\partial 1) \ln \frac{L}{\epsilon}$ is absent in odd spacetime dimensions.

The UV-divergence of the entanglement entropy is one of our main motivations to study *relative entanglement entropies* in Chapter 7, which do not exhibit such divergences and are well-defined in the continuum limit $\epsilon \to 0$ [193, 208, 209] (see also [210–224] for recent works).

EXAMPLE: ENTANGLEMENT ENTROPY FOR THERMAL STATE We exemplify the different contributions in the area law (4.22) for a thermal state with temperature T in a conformal QFT in $d = 1 + 1$ spacetime dimensions. A conformal QFT is invariant under conformal transformations and hence no other scales are present. We expect from (4.22) a logarithmic divergence in leading order. We consider an interval 1 of length L with infinite complement 2. In $d = 1 + 1$ dimensions, the volume of the interval is given by its length $V = L$. We find for the entanglement entropy [195]

$$S(\rho_1) = -\frac{c}{3} \ln \epsilon + \frac{c}{3} \ln \frac{\sinh \pi T L}{\pi T} + \kappa, \tag{4.23}$$

where c denotes the central charge of the conformal field theory and κ is some non-universal constant. Besides the universal UV-divergence from the entangling surface $\partial 1$, we obtain a finite part $\tilde{S}$ depending on the interval length L and the temperature T.

As the length L and the inverse temperature $1\beta = \frac{1}{T}$ provide two competing length scales, let us consider the two asymptotic limits of small and large thermal correlations lengths compared to the interval size

$$\tilde{S} = \begin{cases} \frac{c}{3} \ln L + \kappa & \text{for} \quad L \ll \beta, \\ \frac{c}{3} \pi T L + \kappa & \text{for} \quad L \gg \beta. \end{cases} \tag{4.24}$$

This shows that the finite part of the entanglement entropy $\tilde{S}$ becomes independent of the temperature T for small temperatures $L \ll \beta$ and reduces to the vacuum result. Moreover, $\tilde{S}$ is extensive for sufficiently large temperatures $L \gg \beta$ and in fact corresponds to the thermal entropy in a conformal QFT of $d = 1 + 1$ spacetime dimensions.

BEYOND THE SCOPE   It is important to note that not all QFTs obey an area law of the form (4.22). For example, in non-relativistic theories with a Fermi surface the area law acquires a logarithmic correction [225, 226]. Also, if the boundary $\partial 1$ is not smooth, for example if it contains cusps, additional UV-divergences arise [227].

Furthermore, it is worth mentioning that the entanglement entropy can exhibit Infrared (IR)-divergences, for example when not including a finite mass m, which typically acts as an IR-regulator, or if the theory lives on a finite domain. These divergences do also carry over to the relative entanglement entropy as well as the quantum mutual information [194].

Part II

EQUILIBRIUM

We investigate how equilibrium can be described in terms relative entropy and thermal *model* states. To that end, we establish a statistical inference method based on relative entropy (Chapter 5). We develop thermodynamics solely with relative entropy and use the monotonicity of relative entropy to derive second-law like inequalities (Chapter 6). These ideas are generalized to relativistic fluids in the context of local quantum field theory (Chapter 7).

5 RELATIVE ENTROPY PRINCIPLES

This chapter is based on the first half of [F]. S. F. proposed and supervised the project. The principle of minimum expected relative entropy was found by me. It was formalized and generalized to quantum states by both authors. While both authors participated in the writing of the text, I wrote the early versions of the draft. Also, I created all figures.

We develop a statistical inference method to find a prior distribution or density operator based on (quantum) relative entropy, which we call the *minimum expected relative entropy principle*. We will show that it corresponds to a reformulation of the maximum entropy principle. Further, we will demonstrate that it allows for an information geometric interpretation of a prior distribution for a simple example.

For relative entropy, there already exists the principle of minimum discrimination information (also known as principle of minimum cross entropy), which we review in Section 5.1 for classical distributions [20, 56]. It is used to *update* a *given* prior distribution based on side information, for example in the context of Monto Carlo simulations or machine learning [228, 229]. In this sense, it is *not* capable of providing a prior distribution, which raises the need for an alternative.

Our main motivation is to apply our methods in the context of statistical physics, more precisely to formulate thermodynamics in terms of relative entropy only, which is discussed in detail in Chapter 6. We discuss classical distributions in Section 5.2 and then generalize the formalism to density operators in Section 5.3. We concentrate on discrete variables with finite alphabets and finite Hilbert spaces.

5.1 UPDATING A PRIOR DISTRIBUTION

UPDATING KNOWLEDGE WITH EXPECTATION VALUES   We consider a system whose discrete microstates x are drawn from a finite and countable alphabet $\mathcal{X}$ of size $|\mathcal{X}| = N$ with $N < \infty$. Let us further assume that we already constructed a discrete prior distribution $\tilde{p}(x)$ for the microstates $x \in \mathcal{X}$, which is in accordance with our knowledge about the system. Afterwards, we perform j additional experiments by measuring other observables A_i, where $i \in \{1, ..., j\}$. Each of these j observables takes values $a_i(x)$ in the microstates $x \in \mathcal{X}$. As this additional information about the system is not contained in our prior $\tilde{p}(x)$, we want to find a new distribution $p(x)$, which takes this side information into account.

To that end, we apply the principle of minimum discrimination information, which states that we find $p(x)$ by minimizing the relative entropy $S(p\|\tilde{p})$ with respect to $p(x)$ under the j additional constraints and normalization of $p(x)$. Intuitively, this means that we search for the distribution $p(x)$ which is *least distinguishable* from $\tilde{p}(x)$ and additionally respects the available side information.

We formulate the optimization problem in terms of j Lagrangian multipliers λ_i, which are assigned to the j expectation values of the observables A_i, and another multiplier γ, which ensures normalization of $p(x)$. Hence, we formulate a Lagrange function

$$\mathcal{L}(p, \tilde{p}, \lambda_i, \gamma)$$
$$= S(p\|\tilde{p}) + \sum_{i=1}^{j} \lambda_i \left(\sum_{x \in \mathcal{X}} p(x) a_i(x) - A_i \right) + \gamma \left(\sum_{x \in \mathcal{X}} p(x) - 1 \right) \tag{5.1}$$

and search for a minimum

$$\delta\mathcal{L}(p, \tilde{p}, \lambda_i, \gamma) = 0 \tag{5.2}$$

with respect to a variation in $p(x)$. This optimization problem can be solved in a straightforward manner and leads to

$$p(x) = \frac{\tilde{p}(x)}{Z'} e^{-\sum_{i=1}^{j} \lambda_i a_i(x)}, \tag{5.3}$$

wherein Z' is a normalization factor. We observe that the updated distribution $p(x)$ is proportional to $\tilde{p}(x)$ with the proportionality constants being weighting factors containing the additional side information about the j observables A_i.

EXAMPLE: SPINS IN MAGNETIC FIELD   To apply the principle of minimum discrimination information to a physical problem, we consider a collection of spins, which add up to a macroscopic spin. The orientations of an individual spin make up the microstates $x \in \mathcal{X}$, to which we associate a magnetic moment $\vec{\mu}(x) = \gamma\vec{s}(x)$. Therein, $\vec{s}(x)$ denote the possible spin directions and γ is a proportionality constant.

At first, we consider a situation where the macroscopic spin is in thermal equilibrium with a heat bath of inverse temperature β and in a ferromagnetic configuration at vanishing *external* magnetic field $H = 0$. Evaluating the principle of maximum entropy gives us the prior distribution

$$\tilde{p}(x) = \frac{1}{\tilde{Z}} e^{-\beta E(x)} \tag{5.4}$$

for the microstates x of the single spin. Without additional assumptions, the individual spin directions $\vec{s}(x)$ are not constrained and hence all directions are equally likely.

If we consider a group of individual spins, it is possible that we measure a non-vanishing magnetization $\vec{M}$ in this region, especially if the

macroscopic spin is in a ferromagnetic state. We can update the prior (5.4) based on the additional side information about such a region, i.e. we further constrain the magnetic moment $\vec{\mu}(x)$ by a given *local* magnet field $\vec{B}$. Applying the principle of minimum discrimination information leads us to the updated distribution

$$p(x) = \frac{\tilde{p}(x)}{Z'} e^{\beta \gamma \vec{B} \vec{s}(x)} = \frac{1}{Z} e^{-\beta (E(x) - \gamma \vec{B} \vec{s}(x))} \qquad (5.5)$$

with normalization

$$Z = \tilde{Z} Z' = \sum_{x \in \mathcal{X}} e^{-\beta (E(x) - \gamma \vec{B} \vec{s}(x))}. \qquad (5.6)$$

In principle, the same logic can be applied for other regions in the spin system as well. In this way, we would obtain a simple mean-field description for the macroscopic spin.

APPLICATION TO THERMODYNAMICS Although the updated distribution (5.3) is of Boltzmann-type, it is not possible to apply the principle of minimum discrimination information to thermodynamics in a straightforward way in general. More precisely, it is tempting to conclude that if we would have started from a microcanonical ensemble, i.e. a uniform distribution with respect to energies of individual microstates, and included a given energy expectation value via (5.3), we would have ended up with a canonical distribution. However, for most systems, a uniform distribution with respect to energies is not a physical situation, but rather a formal limit of the canonical ensemble with infinite temperature. As we will see later, the microcanonical ensemble does rather assume a fixed total energy and asserts equal probabilities to the microstates which are compatible with this assumption.

To summarize, the principle of minimum discrimination information is neither capable of predicting a prior distribution nor can it be applied easily to thermodynamics. Since we are interested in developing thermodynamics from a relative entropy perspective, we will develop a new principle of inference in the following.

5.2 FINDING A PRIOR CLASSICAL DISTRIBUTION

INTEGRATING ON THE SPACE OF DISTRIBUTIONS Let us consider again a discrete random variable $x \in \mathcal{X}$ with countable $|\mathcal{X}| = N < \infty$. We want to construct an integral measure on the space of probability distributions $p(x)$. To that end, we have to respect all available constraints on $p(x)$, such that the integration is only over those distributions, which are compatible with these constraints. As always, a minimal constraint is given by the normalization of distributions, i.e. Kolmogorov's second axiom (2.3). However, as we saw in many previous examples, one often has additional constraints, which nevertheless allow for a large class of compatible probability distributions.

To formalize this idea, it is important to note that the set of all probability distributions corresponds to a manifold. Hence, we can find coordinates $\xi = \{\xi^1, ..., \xi^m\}$ which parameterize the set of allowed distributions, i.e. $p(x) = p(x, \xi)$. Then, we can define an integral measure on this manifold via

$$\int \mathcal{D}p = \int d\xi^1 ... d\xi^m \, \mu(\xi^1, ..., \xi^m), \tag{5.7}$$

wherein $\mu(\xi^1, ..., \xi^m)$ is a function to be defined. Note that we have chosen the functional integral notation in order to be as general as possible. More precisely, if we consider continuous random variables instead, the space of distributions becomes infinite-dimensional in general, which requires an integration over functions.

DIFFEOMORPHISM INVARIANCE To determine $\mu(\xi^1, ..., \xi^m)$, we require that our integral measure $\mathcal{D}p$ is invariant under an arbitrary coordinate transformation $\xi \to \xi'(\xi)$ resulting in

$$\mu(\xi^1, ..., \xi^m) = \det\left(\frac{\partial \xi'^\alpha}{\partial \xi^\beta}\right) \mu'(\xi'^1, ..., \xi'^m). \tag{5.8}$$

This is fulfilled for example by Jeffreys' prior [230, 231]

$$\mu(\xi) \propto \sqrt{\det g_{\alpha\beta}(\xi)}, \tag{5.9}$$

wherein $g_{\alpha\beta}(\xi)$ is the so-called Fisher metric corresponding to the family of distributions $p(x, \xi)$ with $\alpha, \beta = 1, ..., m$. Its components are defined as (there are two other convenient definitions in the literature, which are equivalent to what follows) [232, 233]

$$g_{\alpha\beta}(\xi) = \sum_{x \in \mathcal{X}} \frac{\partial p(x, \xi)}{\partial \xi^\alpha} \frac{\partial \ln p(x, \xi)}{\partial \xi^\beta}. \tag{5.10}$$

Note that for a continuous random variable x, the sum in the latter expression is replaced by an integral. One may also introduce the Fisher metric as the Hessian matrix of the relative entropy, i.e. for close distributions $p(x, \xi)$ and $p(x, \xi + d\xi)$ we find

$$S(p(x, \xi + d\xi) \| p(x, \xi)) = \frac{1}{2} g_{\alpha\beta}(\xi) d\xi^\alpha d\xi^\beta + ... \tag{5.11}$$

Moreover, the Fisher metric is symmetric $g_{\alpha\beta}(\xi) = g_{\beta\alpha}(\xi)$ and corresponds to the Riemannian metric on the space of probability distributions. Hence, the canonical volume form $\sqrt{\det g_{\alpha\beta}(\xi)}$ ensures diffeomorphism invariance of our integral measure $\mathcal{D}p$. Nevertheless, we could multiply $\mathcal{D}p$ with any other diffeomorphism invariant function. For example, one may consider $e^{-S(p\|\tilde{p})}$ with $\tilde{p}(x)$ being some model distribution. However, we work with the volume form in the following by convenience. In conclusion, we end up with a measure which allows us to integrate a functional $f[p]$ of the distribution $p(x)$

$$\int \mathcal{D}p \, f[p] = \int d\xi^1 ... d\xi^m \sqrt{\det g_{\alpha\beta}(\xi)} \, f[p(\xi)]. \tag{5.12}$$

PERMUTATION INVARIANCE   Let us analyze this integration proce-
dure in more detail. In particular, we ask how the integral measure
$\int \mathcal{D}p$ transforms if we consider permutations of the individual discrete
probabilities, under which the entropy $S(p)$ is known to be invariant.
We denote a permutation by $x \to \Pi(x)$, such that a permutation on
the probability distribution $p(x)$ is given by the map

$$p \to \Pi(p) : \{p(x_1), ..., p(x_N)\} \to \{p(\Pi(x_1)), ..., p(\Pi(x_N))\}. \quad (5.13)$$

To answer this question, we choose a specific set of coordinates. In
information geometry, it is convenient to parameterize the discrete
probabilities $p(x)$ in a way such that we are confined to a sphere. This
can be achieved by the choice [13, 233, 234]

$$p(x) = \begin{cases} (\xi^x)^2 & \text{for } x = 1, ..., N-1, \\ 1 - (\xi^1)^2 - ... - (\xi^{N-1})^2 & \text{for } x = N. \end{cases} \quad (5.14)$$

In this coordinates, the Fisher metric takes the form

$$\frac{1}{4} g_{\alpha\beta}(\xi) = \delta_{\alpha\beta} + \frac{\xi^\alpha \xi^\beta}{1 - (\xi^1)^2 - ... - (\xi^{N-1})^2}, \quad (5.15)$$

which is the induced metric on the unit sphere S^{N-1}. Geometrically, the
probabilities $p(x)$ span a convex simplex embedded in N-dimensional
Euclidean space. Points on this simplex are mapped to a $(N-1)$-
dimensional sphere, which is described by the coordinates ξ^α. More-
over, the second line in (5.14) shows that the spherical geometry in
the coordinates ξ^α is a consequence of the normalization condition
$\sum_{x \in \mathcal{X}} p(x) = 1$.

In the ξ coordinates, our integral measure $\mathcal{D}p$ takes the form

$$\int \mathcal{D}p = \frac{2}{\Omega_N} \int_{-1}^{1} d\xi^1 ... d\xi^{N-1} \sqrt{\det\left(\frac{g_{\alpha\beta}(\xi)}{4}\right)} \Theta\left(1 - \sum_{\alpha=1}^{N-1} (\xi^\alpha)^2\right)$$

$$= \frac{1}{\Omega_N} \int_{-1}^{1} d\xi^1 .. d\xi^N \, \delta\left(1 - \sqrt{\sum_{\alpha=1}^{N} (\xi^\alpha)^2}\right), \quad (5.16)$$

where

$$\Omega_N = \frac{2\pi^{\frac{N}{2}}}{\Gamma\left(\frac{N}{2}\right)} \quad (5.17)$$

denotes the area element of the $(N-1)$-dimensional unit sphere with
$\Gamma(t)$ being the Gamma function. From the latter equation, we can draw
two important conclusions. First, the integral measure $\mathcal{D}p$ is indeed
normalized to one $\int \mathcal{D}p = 1$ and second, it is invariant under the
permutations $p \to \Pi(p)$ introduced in (5.13). Hence, we end up with
the statement

$$\int \mathcal{D}p \, f[p] = \int \mathcal{D}p \, f[\Pi(p)] \quad (5.18)$$

for all functionals $f[p]$ and permutations Π, which will be of great use
in the following.

CLASSICAL PRINCIPLE OF MINIMUM EXPECTED RELATIVE ENTROPY
With a suitable integral measure $\mathcal{D}p$ at hand, we now state our principle of minimum expected relative entropy or principle of least distinguishability on average. The basic idea is that the optimal prior $\tilde{p}(x)$ is located at a *central position* in the space of probability distributions (see also dark brown point in Figure 3.2). This central point is determined by minimizing the average relative entropy $S(p\|\tilde{p})$ between all allowed distributions $p(x)$ and the optimal model $\tilde{p}(x)$ under the constraint that $\tilde{p}(x)$ is normalized to one. Hence, we construct the functional

The optimal model distribution is least distinguishable on average from all other possible distributions.

$$B(\tilde{p}, \lambda) = \int \mathcal{D}p \left[S(p\|\tilde{p}) + \lambda \left(\sum_{x \in \mathcal{X}} \tilde{p}(x) - 1 \right) \right] \quad (5.19)$$

with normalization being implemented via the Lagrangian multiplier λ and claim that the optimal model $\tilde{p}(x)$ follows from minimizing the latter with respect to $\tilde{p}(x)$, i.e.

$$\delta B(\tilde{p}, \lambda) \stackrel{!}{=} 0. \quad (5.20)$$

Strict joint convexity of relative entropy $S(p\|\tilde{p})$ together with monotonicity of the integral imply that the B-functional is also strictly convex. Hence, if a local minimum of $B(\tilde{p})$ can be found, it corresponds to a unique global minimum.

To obtain the optimal model $\tilde{p}(x)$, we directly compute the variation under the integral (instead of evaluating the integral first), which leads us to

$$0 = \delta B(\tilde{p}, \lambda) = \sum_{x \in \mathcal{X}} \int \mathcal{D}p \left[-\frac{p(x)}{\tilde{p}(x)} + \lambda \right] \delta \tilde{p}(x). \quad (5.21)$$

As a consequence of permutation invariance (5.18), we can conclude that $\int \mathcal{D}p\, p(x)$ is independent of x and hence, $\tilde{p}(x)$ is also independent of x. By normalization, this implies that the optimal model is given by the uniform distribution

$$\tilde{p}(x) = \frac{1}{N}, \quad (5.22)$$

in accordance with Jaynes' maximum entropy principle and Laplace's principle of indifference.

EXAMPLE: TWO–STATE SYSTEM – PART I   To shape the intuition for the principle of minimum expected relative entropy, we consider a coin with $\mathcal{X} = \{\text{heads}, \text{tails}\}$ such that $|\mathcal{X}| = 2$ and compute all involved quantities. The parameterization given in (5.14) evaluates to

$$p(\text{heads}) \equiv p_1 = (\xi^1)^2, \quad p(\text{tails}) = 1 - p_1 = 1 - (\xi^1)^2, \quad (5.23)$$

such that the Fisher information (which is now a scalar) becomes

$$\frac{1}{4}g = 1 + \frac{(\xi^1)^2}{1 - (\xi^1)^2}. \quad (5.24)$$

One can convince oneself that the integration measure $\mathcal{D}p$ is normalized to one as

$$\int \mathcal{D}p = \frac{2}{\Omega_2} \int_{-1}^{1} \mathrm{d}\xi^1 \sqrt{1 + \frac{(\xi^1)^2}{1 - (\xi^1)^2}} = \frac{2}{2\pi} \pi = 1, \qquad (5.25)$$

where we used that the Θ-function in (5.16) evaluates to unity. Similarly, we find for the two individual probabilities

$$\int \mathcal{D}p \, p_1 = \int \mathcal{D}p \, (1 - p_1) = \frac{1}{2} \qquad (5.26)$$

and for the B-functional at the extremum with respect to λ

$$\begin{aligned}
B(\tilde{p}) &= \int \mathcal{D}p \, S(p\|\tilde{p}) \\
&= -\int \mathcal{D}p \, S(p) - \sum_{x \in \mathcal{X}} \ln \tilde{p}(x) \int \mathcal{D}p \, p(x) \qquad (5.27) \\
&= 1 - 2\ln 2 - \frac{1}{2} \ln\left(\tilde{p}_1 (1 - \tilde{p}_1)\right),
\end{aligned}$$

where we used $\tilde{p}(\text{heads}) \equiv \tilde{p}_1$ such that $\tilde{p}(\text{tails}) = 1 - \tilde{p}_1$.

The first line of the latter equation suggests that the value of the B-functional at a specific point $\tilde{p}_1$ can be regarded as the area under a relative entropy curve $S(p\|\tilde{p})$ for a fixed model $\tilde{p}_1$. We make this connection explicit by plotting the relative entropy $S(p\|\tilde{p})$ as a function of p_1 for three different models $\tilde{p}_1$ in Figure 5.1a and the B-functional $B(\tilde{p})$ together with the entropy $S(\tilde{p})$ as a function of $\tilde{p}_1$ in Figure 5.1b.

(a) Three relative entropies $S(p\|\tilde{p})$ (b) B-functional $B(\tilde{p})$ and entropy $S(\tilde{p})$
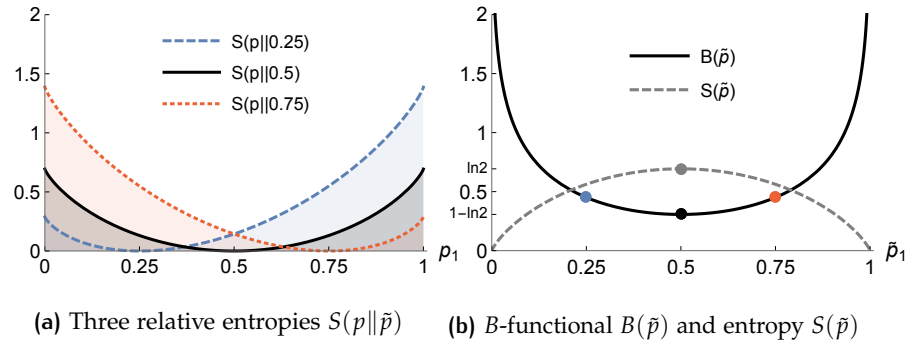
Figure 5.1: In (a), we observe that the relative entropy $S(p\|\tilde{p})$ of a coin with respect to the optimal model $\tilde{p}_1 = 0.5$ takes its minimum at $p_1 = 0.5$ and is symmetric. For other models, $S(p\|\tilde{p})$ is asymmetric and becomes minimal at $p_1 = \tilde{p}_1$. Additionally, we plot $B(\tilde{p})$ and $S(\tilde{p})$ in (b). Interestingly, $B(\tilde{p})$ corresponds to the area under a relative entropy curve $S(p\|\tilde{p})$ (shaded regions in (a) correspond to dots in (b)), such that it indeed becomes minimal for $\tilde{p}_1 = 0.5$. At this point, the entropy $S(\tilde{p})$ attains its maximum (gray dot), showing consistency between the two entropy principles. Also, we see that the B-functional is strictly convex rendering the local minimum at $\tilde{p}_1 = 0.5$ a unique global minimum.

INFORMATION GEOMETRIC INTERPRETATION WITH HELLINGER DISTANCE   Although our principle of least distinguishability on average was formulated with relative entropy $S(p\|\tilde{p})$ as a measure for the distinguishability between two distributions $p(x)$ and $\tilde{p}(x)$, one should be able to replace the relative entropy $S(p\|\tilde{p})$ by any other statistical divergence, i.e. a non-negative (possibly convex) functional being zero if and only if $p(x) = \tilde{p}(x)$, which defines a Riemannian metric in second order. We collect some evidence in the following, but a formal proof of this claim is left for future work.

To give our principle a solid information geometric interpretation, we consider a special choice for the statistical divergence: the Hellinger distance $H(p, \tilde{p})$. It is proportional the Euclidean distance on the space of probability distributions [235, 236]

$$H(p, \tilde{p}) = \frac{1}{\sqrt{2}} \left\| \sqrt{p} - \sqrt{\tilde{p}} \right\|_2 = \frac{1}{\sqrt{2}} \sqrt{\sum_{x \in \mathcal{X}} \left(\sqrt{p(x)} - \sqrt{\tilde{p}(x)} \right)^2}, \quad (5.28)$$

where $\|.\|_2$ denotes the L_2-norm, which equals the Euclidean distance. Note that the prefactor $\frac{1}{\sqrt{2}}$ ensures that the Hellinger distance is bounded from above by one, i.e. $0 \leq H(p, \tilde{p}) \leq 1$, which is a consequence of the Cauchy-Schwarz inequality. Also, note the relation to the relative entropy $2H^4(p, \tilde{p}) \leq S(p\|\tilde{p})$.

The Hellinger distance $H(p, \tilde{p})$ is a statistical divergence. Additionally, it is symmetric and fulfills a triangle inequality. Therefore, it is a *true* distance measure on the space of probability distribution, which allows us to make more rigorous statements when it comes to information geometry.

With the Hellinger distance $H(p, \tilde{p})$, our B-functional reads

$$B(\tilde{p}, \lambda) = \int \mathcal{D}p \left[H(p, \tilde{p}) + \lambda \left(\sum_{x \in \mathcal{X}} \tilde{p}(x) - 1 \right) \right] \quad (5.29)$$

and the optimal model $\tilde{p}(x)$ still follows from the variational principle $\delta B(\tilde{p}, \lambda) = 0$ with a variation with respect to $\tilde{p}(x)$.

EXAMPLE: TWO-STATE SYSTEM – PART II   We will not compute the variation under the integral as done before, but rather evaluate the B-functional for the simple coin example. Unfortunately, the complicated structure of the integrand does only allow for a numerical evaluation.

We show Hellinger distance curves in Figure 5.2a and the Hellinger B-functional in Figure 5.2b. Most importantly, our intuition from the formulation in terms of relative entropies can be confirmed: the optimal model $\tilde{p}(x)$ has the least mean Hellinger distance to all other allowed distributions $p(x)$. Hence, the optimal model $\tilde{p}(x)$ is indeed the *central point* on the manifold of distributions. Exemplary, we also compare a collection of Hellinger distances for two models in Figure 5.2c, showing once again that the uniform model is preferred.
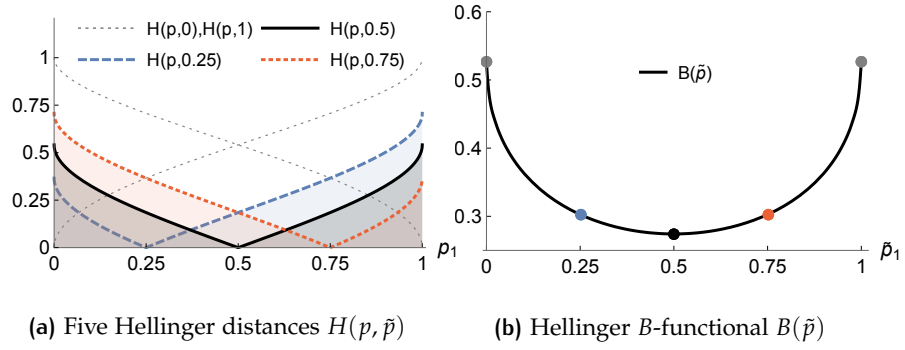
(a) Five Hellinger distances $H(p, \tilde{p})$

(b) Hellinger B-functional $B(\tilde{p})$
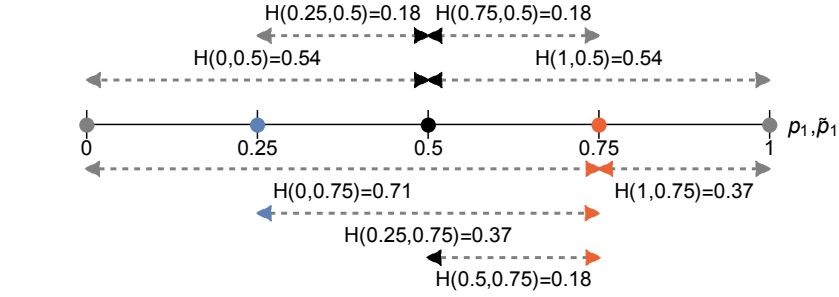
(c) Exemplary Hellinger distances for the two models $\tilde{p}_1 = 0.5, 0.75$

Figure 5.2: (a) and (b) are similar to Figure 5.1 with the relative entropy $S(p\|\tilde{p})$ being replaced by the Hellinger distance $H(p, \tilde{p})$. We have added two curves for the two certain models $\tilde{p}_1 = 0$ and $\tilde{p}_1 = 1$ in (a) showing that the Hellinger distance $H(p, \tilde{p})$ is bounded from above by one. Note the relative entropies $S(p\|0)$ and $S(p\|1)$ are either zero or infinity in this case. As a consequence, the Hellinger B-functional is also finite at $\tilde{p}_1 = 0$ and $\tilde{p}_1 = 1$, which is emphasized in (b) through the two additional gray dots. In (c), we show some exemplary Hellinger distances for the optimal model $\tilde{p}_1 = 0.5$ and a second model $\tilde{p}_1 = 0.75$. One can easily confirm that the mean Hellinger distance in the upper half is smaller than in the lower half, although the mean has not been taken with respect to all allowed distributions p_1.

A NOTE ON OTHER CONSTRAINTS   If other constraints are given, the situation becomes much more involved as they have to be respected for the integration over all allowed distributions $p(x)$ as well as for the optimal model $\tilde{p}(x)$. For the latter, adding Lagrange multipliers suffices, while for the former we have to adjust our integral measure $\mathcal{D}p$. However, a similar parameterization as in (5.14) is typically not favorable, as the integration domain does no longer correspond to the surface of a sphere. Nevertheless, from (5.21) together with permutation invariance (5.18) we can conclude that the optimal model should in any case correspond to the expectation value with respect to all allowed $p(x)$, i.e. $\tilde{p}(x) = \langle p(x) \rangle$, where $\langle . \rangle = \int \mathcal{D}p(.)$. However, a more detailed investigation, especially for concrete examples, is left for the future.

A NOTE ON CONTINUOUS RANDOM VARIABLES   Another interesting route for future work concerns continuous random variables. In this case, our integral measure $\mathcal{D}p$ becomes a functional integral over probability density functions $f(x)$ which are normalized to unity. Therefore, the corresponding space of allowed distributions is infinite-dimensional im general. It would be of great interest to see whether the principle of minimum expected relative entropy is capable of predicting the uniform model $\tilde{f}(x) = 1/(b-a)$ when restricting the random variable to a finite interval $x \in [a, b]$. If yes, one may also attempt to implement further constraints, for example a given variance σ_x^2.

5.3 FINDING A PRIOR QUANTUM STATE

INTEGRATING ON THE SPACE OF DENSITY OPERATORS   We now generalize our considerations to density operators ρ which act on a Hilbert space $\mathcal{H}$ of finite dimension $D < \infty$. We are interested in constructing an integral measure $\mathcal{D}\rho$, which allows us to integrate over the space of density operators ρ. In complete analogy to the classical case we choose coordinates $\xi = (\xi^1, ..., \xi^m)$ to parameterize this space. Also, we again use Jeffreys' prior for the canonical volume form, such that the integral measure $\mathcal{D}\rho$ is by construction invariant under arbitrary coordinate transformations $\xi \to \xi'(\xi)$. Then, a functional $f[\rho]$ can be integrated according to

$$\int \mathcal{D}\rho\, f[\rho] = \int \mathrm{d}\xi^1 ... \mathrm{d}\xi^m\, \sqrt{\det g_{\alpha\beta}(\xi)}\, f[\rho(\xi)]. \tag{5.30}$$

Therein, we used the *quantum* Fisher metric instead of its classical counterpart, which is defined as [13, 237, 238] (see also [239–246] for recent applications)

$$g_{\alpha\beta}(\xi) = \mathrm{Tr}\left\{ \frac{\partial\rho(\xi)}{\partial\xi^\alpha} \frac{\partial \ln\rho(\xi)}{\partial\xi^\beta} \right\}, \tag{5.31}$$

with the following definition for the symmetric logarithmic derivative understood

$$\frac{1}{2}\rho(\mathrm{d}\ln\rho) + \frac{1}{2}(\mathrm{d}\ln\rho)\rho = \mathrm{d}\rho. \tag{5.32}$$

Here, note that $\mathrm{Tr}\{\rho(\mathrm{d}\ln\rho)\} = \mathrm{Tr}\{\mathrm{d}\rho\} = 0$, which implies that also the quantum Fisher metric is symmetric $g_{\alpha\beta}(\xi) = g_{\beta\alpha}(\xi)$. Furthermore, the quantum Fisher metric corresponds to the Hessian matrix of the quantum relative entropy, i.e.

$$S(\rho(\xi + \mathrm{d}\xi) \| \rho(\xi)) = \frac{1}{2}g_{\alpha\beta}(\xi)\mathrm{d}\xi^\alpha \mathrm{d}\xi^\beta + ... \tag{5.33}$$

UNITARY INVARIANCE Instead of permutations as for the classical case, we are now interested in unitary transformations $\rho \to U\rho U^\dagger$, which leave the von Neumann entropy invariant (cf. (3.16)). As the transformed density operator is itself a density operator again, a unitary transformation corresponds to a coordinate transformation $\xi \to \xi'(\xi)$, i.e.

$$\rho(\xi) \to \rho'(\xi) = U\rho(\xi)U^\dagger = \rho(\xi'). \tag{5.34}$$

The coordinate transformation induced by the unitary operator U can be constrained further by plugging the latter equation into the relative entropy of nearby states

$$\begin{aligned} S(\rho(\xi + \mathrm{d}\xi)\|\rho(\xi)) &= S(U\rho(\xi + \mathrm{d}\xi)U^\dagger\|U\rho(\xi)U^\dagger), \\ &= S(\rho(\xi' + \mathrm{d}\xi')\|\rho(\xi')) \end{aligned} \tag{5.35}$$

where we used that the quantum relative entropy is also left invariant by any unitary transformation. Comparing with (5.33) shows

$$g_{\alpha\beta}(\xi)\mathrm{d}\xi^\alpha\mathrm{d}\xi^\beta = g_{\alpha\beta}(\xi')\mathrm{d}\xi'^\alpha\mathrm{d}\xi'^\beta, \tag{5.36}$$

implying that $\xi \to \xi'(\xi)$ in fact corresponds to an isometry on the space of density operators. Hence, we can conclude that our integral measure $\mathcal{D}\rho$ behaves as

$$\int \mathcal{D}\rho\, f[\rho] = \int \mathcal{D}\rho\, f[U\rho U^\dagger] \tag{5.37}$$

for any functional $f[\rho]$ under any unitary transformation U.

QUANTUM PRINCIPLE OF MINIMUM EXPECTED RELATIVE ENTROPY
The optimal model state $\tilde{\rho}$ in a class of allowed states ρ can be found by applying the quantum analog of the classical principle of minimum expected relative entropy (5.20). To that end, we define the quantum B-functional as an average of the quantum relative entropy $S(\rho\|\tilde{\rho})$

$$B(\tilde{\rho}, \lambda) = \int \mathcal{D}\rho\, [S(\rho\|\tilde{\rho}) + \lambda\,(\mathrm{Tr}\{\tilde{\rho}\} - 1)], \tag{5.38}$$

The optimal model state is least distinguishable on average from all other possible states.

where λ ensures normalization of $\tilde{\rho}$. Then, we claim that the optimal prior $\tilde{\rho}$ is the result of the minimization procedure

$$\delta B(\tilde{\rho}, \lambda) \stackrel{!}{=} 0. \tag{5.39}$$

To solve this variational principle, we diagonalize the model state $\tilde{\rho}$ with the help of a unitary transformation U

$$\tilde{\rho} = U\Delta U^\dagger, \tag{5.40}$$

where Δ is diagonal. Then, the quantum B-functional becomes

$$\begin{aligned} B(\tilde{\rho}) &= \int \mathcal{D}\rho\, \left[S(\rho\|U\Delta U^\dagger) + \lambda\left(\mathrm{Tr}\left\{U\Delta U^\dagger\right\}\right) - 1\right] \\ &= \int \mathcal{D}\rho\, \left[S(U^\dagger\rho U\|\Delta) + \lambda\left(\mathrm{Tr}\left\{\Delta\right\}\right) - 1\right] \\ &= \int \mathcal{D}\rho\, \left[S(\rho\|\Delta) + \lambda\left(\sum_{i=1}^{D}\Delta_{ii} - 1\right)\right], \end{aligned} \tag{5.41}$$

where we used again that the quantum relative entropy is invariant under unitary transformations and that the integral measure fulfills the invariance property (5.37). Also, we denoted the diagonal elements of Δ in some generic basis as Δ_{ii}. In conclusion, the quantum B-functional does only depend on the eigenvalues of $\tilde{\rho}$ and gives the same result for all model states $\tilde{\rho}$ which are related by unitary transformations.

For simplicity, we directly execute the variation with respect to the diagonal elements Δ_{ii}, which leads to

$$0 = \delta B(\tilde{\rho}, \lambda) = \sum_{i=1}^{D} \int \mathcal{D}\rho \left[-\frac{\rho_{ii}}{\Delta_{ii}} + \lambda \right] \delta\Delta_{ii}, \tag{5.42}$$

in analogy to (5.21). The diagonal elements ρ_{ii} can be permuted by special unitary transformations (for example, for $D = 2$ one may use the second Pauli operator $U = \sigma_2$), i.e. $\rho_{ii} \to \rho_{\Pi(i)\Pi(i)}$ for $\rho \to U\rho U^\dagger$ with suitable U, such that the unitary invariance property (5.37) allows us to conclude that $\int \mathcal{D}\rho\, \rho_{ii}$ is independent of the index i. Hence, the optimal prior is given by the maximally mixed state

$$\tilde{\rho} = \frac{1}{D}\mathbb{1}. \tag{5.43}$$

Interestingly, this result can also be obtained from the unitary invariance property (5.37). As the optimal model $\tilde{\rho}$ has to be unique by strict convexity of the B-functional, (5.41) implies that the optimal model $\tilde{\rho}$ has to satisfy

$$\tilde{\rho} = U\tilde{\rho}U^\dagger \tag{5.44}$$

for all unitary transformations U, which already fully determines $\tilde{\rho}$ to be the normalized identity operator.

A NOTE ON OTHER CONSTRAINTS   Similarly to the classical case, let us briefly comment on the possibility of implementing further constraints. As before, we argue that the optimal prior fulfills $\tilde{\rho} = \langle \rho \rangle$, although any explicit computation of the B-functional becomes complicated as a result of a rather complex integral measure $\mathcal{D}\rho$. To that end, we write out the variation of the B-functional at the extremum with respect to λ, i.e.

$$\delta B(\tilde{\rho}) = \int \mathcal{D}\rho\, \mathrm{Tr}\left\{\rho(\mathrm{d}\ln\tilde{\rho})\right\} \equiv \mathrm{Tr}\left\{\langle\rho\rangle(\mathrm{d}\ln\tilde{\rho})\right\}. \tag{5.45}$$

Using cyclicity of the trace we obtain

$$\delta B(\tilde{\rho}) = \frac{1}{2}\mathrm{Tr}\left\{\langle\rho\rangle(\mathrm{d}\ln\tilde{\rho}) + (\mathrm{d}\ln\tilde{\rho})\langle\rho\rangle\right\} \tag{5.46}$$

For $\tilde{\rho} = \langle\rho\rangle$ and with the definition of the logarithmic derivative (5.32) follows

$$\delta B(\tilde{\rho}) = \mathrm{Tr}\{\mathrm{d}\tilde{\rho}\} = 0, \tag{5.47}$$

as a consequence of $\tilde{\rho}$ being normalized to unity. Hence, the B-functional is indeed stationary for $\tilde{\rho} = \langle\rho\rangle$.

6 | THERMODYNAMICS FROM RELATIVE ENTROPY

Most of the following analysis is taken from the second half of [F]. While the project was proposed and supervised by S. F., I derived the thermodynamic relations in terms of relative entropies and found the relative entropy formulation for the third law. Also, I wrote early versions of the text, which was finalized by both authors. The discussion of the second law is taken from section III. of [G], which is itself based on [247]. The project was proposed by S. F. and the corresponding discussion in [G] was mainly developed by N. D. and myself. All three authors contributed to the writing and structuring of the text, while I created the figures. The article [G] is also covered in a master thesis by N. D. [248].

With the principle of minimum expected relative entropy at hand, we formulate thermodynamics in terms of relative entropy only, for which our motivation is threefold.

First, a *relative* formulation allows to work with a thermal *model* state $\tilde{\rho}$ for given situation. As exemplified before, the principle of minimum expected relative entropy gives these states an information geometric meaning in the space of states. Using a distinguishability measure with respect to these states allows us to relax the necessity of the actual state ρ being thermal. In this sense, relative entropy thermodynamics allows to make statements about thermal quantities as for example temperature and chemical potential by *comparing* actual states to thermal model states.

Second, as stressed throughout Part i, relative entropy may be favored over entropy in general as it has universal properties in various contexts and fulfills the powerful monotonicity relation (3.18). For example, we will show that thermal fluctuations (see [249–251]) can be characterized in terms of relative entropy and that the second law of thermodynamics can be deduced from an information theoretic perspective [247]. Further, relative entropy has been used in the context of thermodynamics in [252–257].

Third, it is of interest to understand how thermodynamic concepts can be applied *locally*, especially in the context of fluids and local QFT. For example, it is known that the quark-gluon plasma [258, 259] or the dark matter dominated cosmological fluid [260, 261] can, to good extent, be described by QFTs on a microscopic level and by relativistic fluids on a macroscopic level. To establish a more general connection between these two classes of theories, an information theoretic understanding of *local* equilibrium may be in order. However, as we have seen in Section 4.2, quantum fields are highly entangled in space,

which resulted in an universal UV-divergence of the entanglement entropy (4.22). Therefore, we suggest to work with relative (entanglement) entropies instead, as this allows for applications beyond standard thermodynamics (see [217]). This will then be investigated further in Chapter 7.

We will discuss thermodynamic relations for the three thermodynamic ensembles in terms of relative entropy in Section 6.1 and thereupon state relative entropic versions of the second and third law of thermodynamics in Section 6.2.

6.1 THERMODYNAMIC ENSEMBLES AND RELATIONS

OVERVIEW OVER THERMODYNAMIC ENSEMBLES Throughout this chapter, we consider the standard setup, i.e. a quantum system confined to a finite volume V described by a stationary Hamiltonian $H \neq H(t)$. For a generic quantum state ρ we denote the expectation values of energy and particle number by

$$E(\rho) = \text{Tr}\{\rho H\}, \quad N(\rho) = \text{Tr}\{\rho N\}. \tag{6.1}$$

Depending on whether the energy or the particle number or their expectation values are fixed, we have to work in one of the three thermodynamic ensembles sketched in Figure 6.1.

(a) microcanonical   (b) canonical   (c) grand canonical
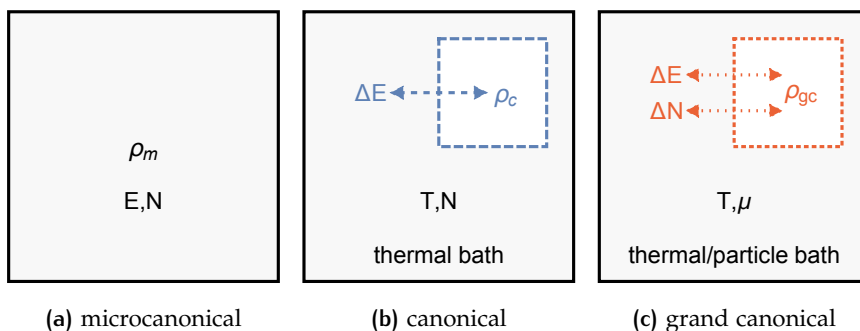
Figure 6.1: The three thermodynamic ensembles
The three thermodynamic ensembles: (a) isolated system with no interaction with the outside, (b) system open to heat exchange in contact with a heat bath and (c) system open to heat and particle exchange in contact with a heat and particle bath.

MICROCANONICAL ENSEMBLE We start with the microcanonical ensemble, i.e. we investigate an isolated quantum system with fixed total energy $H = E(\rho)$ and total particle number $N = N(\rho)$ (cf. Figure 6.1a). In the space of compatible quantum states the uniform distribution is preferred as a result of the principle of minimum expected relative

entropy as shown in Section 5.3. Hence, the microcanonical model state is given by

$$\boldsymbol{\rho}_{\mathrm{m}} = \frac{1}{Z_{\mathrm{m}}} \, \delta\left[\boldsymbol{H} - E(\boldsymbol{\rho}_{\mathrm{m}})\right] \delta\left[\boldsymbol{N} - N(\boldsymbol{\rho}_{\mathrm{m}})\right], \tag{6.2}$$

with Z_{m} being a normalization constant

$$Z_{\mathrm{m}} = \mathrm{Tr}\left\{\delta\left[\boldsymbol{H} - E(\boldsymbol{\rho}_{\mathrm{m}})\right] \delta\left[\boldsymbol{N} - N(\boldsymbol{\rho}_{\mathrm{m}})\right]\right\}. \tag{6.3}$$

The quantum relative entropy between any allowed quantum state $\boldsymbol{\rho}$ and the microcanonical model evaluates to

$$S(\boldsymbol{\rho}\|\boldsymbol{\rho}_{\mathrm{m}}) = -S(\boldsymbol{\rho}) - \mathrm{Tr}\{\boldsymbol{\rho}\ln\boldsymbol{\rho}_{\mathrm{m}}\}, \tag{6.4}$$

wherein we recognize the von Neumann entropy $S(\boldsymbol{\rho})$, while the second term corresponds to the so-called cross entropy often denoted as $S(\boldsymbol{\rho}, \boldsymbol{\rho}_{\mathrm{m}})$. The latter can be simplified by using that $\boldsymbol{\rho}_{\mathrm{m}}$ is proportional to the identity in the space of allowed states

$$-\mathrm{Tr}\{\boldsymbol{\rho}\ln\boldsymbol{\rho}_{\mathrm{m}}\} = -\mathrm{Tr}\{\boldsymbol{\rho}_{\mathrm{m}}\ln\boldsymbol{\rho}_{\mathrm{m}}\} = S(\boldsymbol{\rho}_{\mathrm{m}}), \tag{6.5}$$

where the support condition $\mathrm{supp}[\boldsymbol{\rho}] \subseteq \mathrm{supp}[\boldsymbol{\rho}_{\mathrm{m}}]$ has to be respected. As the microcanonical density operator (6.2) has full support for fixed energy and particle number, the support condition can be expressed as a condition on the total energy and the total particle number. It is fulfilled if and only if the states $\boldsymbol{\rho}$ and $\boldsymbol{\rho}_{\mathrm{m}}$ have the same total energy and total particle number in a *strict* sense. More precisely, not only the expectation values of the latter have to agree, but also the dispersion has to be zero. For the energy E, we denote this strict condition by

$$E(\boldsymbol{\rho}) \equiv E(\boldsymbol{\rho}_{\mathrm{m}}) \Leftrightarrow E(\boldsymbol{\rho}) = E(\boldsymbol{\rho}_{\mathrm{m}}) \text{ and } \mathrm{Tr}\{\boldsymbol{\rho}\boldsymbol{H}^2\} = E^2(\boldsymbol{\rho}_{\mathrm{m}}), \tag{6.6}$$

which works analogously for the particle number N.

The quantum relative entropy with respect to a thermodynamic model state decomposes linearly.

Therefore, we end up with a quantum relative entropy

$$S(\boldsymbol{\rho}\|\boldsymbol{\rho}_{\mathrm{m}}) = \begin{cases} -S(\boldsymbol{\rho}) + S(\boldsymbol{\rho}_{\mathrm{m}}) & \text{for } E(\boldsymbol{\rho}) \equiv E(\boldsymbol{\rho}_{\mathrm{m}}) \\ & \text{and } N(\boldsymbol{\rho}) \equiv N(\boldsymbol{\rho}_{\mathrm{m}}), \\ +\infty & \text{else.} \end{cases} \tag{6.7}$$

This means that the distinguishability between any state $\boldsymbol{\rho}$ and the microcanonical model $\boldsymbol{\rho}_{\mathrm{m}}$ is given by the difference of their von Neumann entropies, provided that their energies and particle numbers agree in a strict sense. Otherwise, the relative entropy has to be set to infinity, which means that the two states can be fully distinguished by measuring either the energy or the particle number.

At this point, note that non-negativity of the relative entropy together with (6.7) shows that the microcanonical model $\boldsymbol{\rho}_{\mathrm{m}}$ does indeed

correspond to the state with maximum entropy as $S(\rho\|\rho_m) \geq 0$ translates to

$$S(\rho) \leq S(\rho_m) \tag{6.8}$$

for all density operators ρ with $E(\rho) \equiv E(\rho_m)$ and $N(\rho) \equiv N(\rho_m)$.

As a next step, let us differentiate the relation (6.7). For $dE(\rho) \equiv dE(\rho_m)$ and $dN(\rho) \equiv dN(\rho_m)$ we find

$$dS(\rho\|\rho_m) = -dS(\rho) + dS(\rho_m). \tag{6.9}$$

Using the thermodynamic relation

$$dS(\rho_m) = \beta\, dE(\rho_m) - \beta\mu\, dN(\rho_m), \tag{6.10}$$

where $\beta = 1/T$ denotes the inverse temperature and μ the chemical potential, the latter becomes

$$dS(\rho\|\rho_m) = -dS(\rho) + \beta\, dE(\rho_m) - \beta\mu\, dN(\rho_m). \tag{6.11}$$

This already indicates that the inverse temperature β as well as the chemical potential μ may be defined as partial derivatives of the relative entropy $S(\rho\|\rho_m)$ with respect to the energy $E(\rho)$ and the particle number $N(\rho_m)$, respectively, at fixed von Neumann entropy $S(\rho)$. This will be made more explicit in the canonical and the grand canonical ensemble.

THERMAL FLUCTUATIONS   We proceed with a discussion of thermal fluctuations from an entropic perspective. It is well-known that even if a system has equilibrated, thermodynamic quantities can fluctuate in subsystems, which can be interpreted as a finite size correction to the thermodynamic limit. Often, those fluctuations are stronger when the subsystem under consideration is smaller, which we attempt to give an information theoretic foundation.

Let us consider some macroscopic variables $\xi = (\xi^1, ..., \xi^m)$, which are allowed to fluctuate locally in a global microcanonical model state ρ_m. To these variables, we may associate a family of density operators $\rho(\xi)$ over a set of microstates, which have to be compatible with the global value of ξ but also with all other possible constraints such as conservation laws. Based on this, we wish to construct an expression for the probability dW of a fluctuation in the volume element $d^m\xi$ in the space of allowed density operators $\rho(\xi)$.

A convenient ansatz in the literature is to take this probability to be proportional to the number of possible microstates, which is given by $e^{S(\rho(\xi))}$ [249, 250]. Note here that $S(\rho(\xi)) < S(\rho_m)$ as a consequence of the maximum entropy principle. For example, one may use $dW \propto e^{S(\rho(\xi))-S(\rho_m)}d^m\xi$, where the exponent resembles the right hand side of (6.7) up to a minus sign.

More generally, we propose that this probability is governed by a relative entropy, i.e. $\mathrm{d}W \propto e^{-S(\rho(\xi)\|\rho_\mathrm{m})}\mathrm{d}^m\xi$. This is also motivated by the fact that $S(\rho(\xi)\|\rho_\mathrm{m})$ is bounded from above by the entropy of the model state $S(\rho_\mathrm{m}) = \ln D_\xi$. Here, $\mathcal{H}_\xi$ denotes the Hilbert space spanned by the microstates being compatible with ξ. Therefore, if the subsystem under consideration is smaller, the Hilbert space dimension D_ξ is typically smaller as well, reducing the upper bound on $S(\rho(\xi)\|\rho_\mathrm{m})$. Hence, states become generically harder to distinguish and as a consequence the probability for local fluctuations in ξ becomes large. Also, if we choose an incompatible state $\rho(\xi)$ in the sense that it does violate the strict support conditions, we obtain a vanishing probability for a fluctuation in ξ, in accordance with our expectations.

Additionally, we want to implement reparameterization invariance, such that we end up with the following expression for the probability to find a fluctuation in ξ

The probability for thermal fluctuations is governed by quantum relative entropy.

$$\mathrm{d}W = \frac{1}{Z} e^{-S(\rho(\xi))\|\rho_\mathrm{m})} \sqrt{\det g_{\alpha\beta}(\xi)}, \tag{6.12}$$

where Z is a normalization constant ensuring $\int \mathrm{d}W = 1$, while $g_{\alpha\beta}(\xi)$ is the quantum Fisher metric with respect to the family of states $\rho(\xi)$.

Let us also consider a rather special case. If we assume that the relative entropy $S(\rho(\xi)\|\rho_\mathrm{m})$ becomes minimal at $\xi = \xi_0$, where we can approximate $\rho(\xi_0) \approx \rho_\mathrm{m}$, we can use that the relative entropy can be expanded in terms of the Fisher metric according to (5.33), corresponding to an approximation quadratic in $\xi - \xi_0$. In this approximation, we obtain

$$\mathrm{d}W = \frac{1}{Z} e^{-\frac{1}{2}g_{\alpha\beta}(\xi)(\xi-\xi_0)^\alpha(\xi-\xi_0)^\beta} \sqrt{\det g_{\alpha\beta}(\xi)}, \tag{6.13}$$

where $Z = (2\pi)^{\frac{m}{2}}$.

We can conclude that fluctuations are to lowest order described by a Gaussian in terms of the Fisher metric $g_{\alpha\beta}(\xi)$ (in agreement with standard literature [250]) and in general by a relative entropy. Let us remark that the previous discussion can be extended to the case where the global model state is given by some other thermodynamic model state by replacing the corresponding model in the relative entropy.

CANONICAL ENSEMBLE   Next, we consider the canonical ensemble describing a quantum system in contact with a heat bath (cf. Figure 6.1b). As discussed in Section 5.3, it is rather hard to show that the canonical state ρ_c can be derived from the principle of minimum expected relative entropy, although it will be a stationary solution in the sense of $\rho_\mathrm{c} = \langle\rho\rangle$. Therefore, we follow the textbook construction of the canonical ensemble, see e.g. [250].

We start from an isolated system described by the microcanonical ensemble with fixed energy $E(\rho) \equiv E(\rho_\mathrm{m})$ and particle number $N(\rho) \equiv N(\rho_\mathrm{m})$ and consider a bipartition 12 of a small subsystem

1 and a heat bath 2 such that the two subsystems are allowed to exchange heat. To the two subsystems, we associate energies $E(\rho_1)$ and $E(\rho_2) = E(\rho_{12}) - E(\rho_1)$, which are not fixed but instead fluctuate as described before.

Although equilbriation requires an interaction between the two subsystems, for the equilibrium state itself we assume the corresponding Hamiltonian to be small compared to the sum of the local Hamiltonians, such that $H_{12} = H_1 + H_2$. Under this assumption, the global microcanonical model state (6.2) can be written as

$$
\rho_m \propto \int_0^{E(\rho_{12})} dE(\rho_1) \frac{dW}{dE(\rho_1)}
\times \delta(H_1 - E(\rho_1))\delta(H_2 - E(\rho_{12}) + E(\rho_1)),
\tag{6.14}
$$

where the integral goes over all local energies $E(\rho_1)$ allowed by energy conservation. Note that the global state ρ_m does *not* correspond to a product state over local states with respect to local energies.

The local states in subsystem 1, which are compatible with energy conservation, can be parameterized by the local energy $\xi = E(\rho_1)$, i.e. $\rho_1 = \rho_1(\xi)$. Let consider a family of product states

$$
\rho_{12}(\xi) = \rho_1(\xi) \otimes \rho_2(E_{12} - \xi),
\tag{6.15}
$$

such that the global relative entropy becomes

$$
S(\rho_{12}(\xi)\|\rho_m) = -S(\rho_1(\xi)) - S(\rho_2(E_{12} - \xi)) + S(\rho_m).
\tag{6.16}
$$

The non-negativity of the relative entropy guarantees the existence of a minimum with respect to ξ, at which we obtain the well-known condition

$$
\frac{\partial}{\partial\xi} S(\rho_1(\xi))\big|_{\xi=E(\rho_1)} = \frac{\partial}{\partial\xi} S(\rho_2(\xi))\big|_{\xi=E(\rho_{12})-E(\rho_1)},
\tag{6.17}
$$

defining the inverse temperature

$$
\beta = \frac{\partial}{\partial\xi} S(\rho_1(\xi))
\tag{6.18}
$$

as an equivalence relation. To determine the state of system 1 in equilibrium, we make use of the expression (6.12) for thermal fluctuations in the variable $\xi = E(\rho_1)$. Expanding the relative entropy given in (6.16) to linear order in β gives

$$
S(\rho_{12}(\xi)\|\rho_m) = \beta E(\rho_1),
\tag{6.19}
$$

which can be inserted in (6.12). Then, replacing the energy $E(\rho_1)$ by the corresponding Hamiltonian H_1 leads to the canonical density operator (we drop the bipartition notation in the following)

$$
\rho_c = \frac{1}{Z_c} e^{-\beta H},
\tag{6.20}
$$

with the canonical partition sum Z_c being given by

$$Z_c = \text{Tr}\{e^{-\beta H}\}. \tag{6.21}$$

Compared to the microcanonical expression (6.2), the energy is not fixed anymore, but rather fluctuates, such that we have a mixture of different energies, which are weighted by Boltzmann-type factors.

The relative entropy between a state ρ and the canonical model ρ_c evaluates to

$$S(\rho\|\rho_c) = -S(\rho) + \ln Z_c + \beta E(\rho), \tag{6.22}$$

when imposing the strict support condition $N(\rho) \equiv N(\rho_c)$. This simplifies further when using

$$S(\rho_c) = \ln Z_c - \beta \frac{\partial}{\partial \beta} \ln Z_c = \ln Z_c + \beta E(\rho_c), \tag{6.23}$$

such that we end up with

$$S(\rho\|\rho_c) = \begin{cases} -S(\rho) + S(\rho_c) + \beta\left[E(\rho) - E(\rho_c)\right] & \text{for } N(\rho) \equiv N(\rho_c), \\ +\infty & \text{else.} \end{cases} \tag{6.24}$$

We find that the relative entropy decomposes linearly into differences of extensive quantities and that the support condition is less strict. More precisely, the relative entropy is finite as long as the particle numbers agree. Also, it is interesting to note that if the energy expectation values agree $E(\rho) = E(\rho_c)$, we get back the result from the microcanonical model (6.7).

As a side note, we report that in [252–255] it was shown that the relative entropy $S(\rho\|\rho_c)$ is proportional to the available energy. If the system is in the state ρ, this corresponds to the maximum work one can extract when coupling the system to a heat bath of inverse temperature β. This can be made more explicit if we choose the state to be thermal itself, i.e. $\rho = \rho'_c$, with some other inverse temperature β'. In this case, one finds that the relative entropy becomes proportional to a difference of (generalized) free energies

$$S(\rho'_c\|\rho_c) = \beta\left[F(\rho_c) - F(\rho'_c)\right], \tag{6.25}$$

where $F(\rho'_c)$ has to be considered as a generalized free energy as it is defined with respect to the inverse temperature β which may differ from β'.

As a next step, we consider the differential of the relative entropy given in (6.24). To that end, we restrict ourselves to the case where the condition $dN(\rho) \equiv dN(\rho_c)$ is fulfilled. We end up with

$$\begin{aligned} dS(\rho\|\rho_c) &= -dS(\rho) + dS(\rho_c) + \left[E(\rho) - E(\rho_c)\right] d\beta \\ &\quad + \beta\left[dE(\rho) - dE(\rho_c)\right] \\ &= -dS(\rho) + \beta\, dE(\rho) - \beta\mu\, dN(\rho) \\ &\quad + \left[E(\rho) - E(\rho_c)\right] d\beta, \end{aligned} \tag{6.26}$$

where we used the thermodynamic relation

$$dS(\boldsymbol{\rho}_c) = \beta \, dE(\boldsymbol{\rho}_c) - \beta\mu \, dN(\boldsymbol{\rho}_c) \tag{6.27}$$

in the second step. For a general quantum state $\boldsymbol{\rho}$, the quantities $S(\boldsymbol{\rho}), E(\boldsymbol{\rho}), N(\boldsymbol{\rho})$ and β are independent. Hence, we find the following four partial differential relations from (6.26)

$$
\begin{aligned}
\left.\frac{\partial S(\boldsymbol{\rho}\|\boldsymbol{\rho}_c)}{\partial S(\boldsymbol{\rho})}\right|_{E(\boldsymbol{\rho}),N(\boldsymbol{\rho}),\beta} &= -1, \\
\left.\frac{\partial S(\boldsymbol{\rho}\|\boldsymbol{\rho}_c)}{\partial E(\boldsymbol{\rho})}\right|_{S(\boldsymbol{\rho}),N(\boldsymbol{\rho}),\beta} &= \beta, \\
\left.\frac{\partial S(\boldsymbol{\rho}\|\boldsymbol{\rho}_c)}{\partial N(\boldsymbol{\rho})}\right|_{S(\boldsymbol{\rho}),E(\boldsymbol{\rho}),\beta} &= -\beta\mu, \\
\left.\frac{\partial S(\boldsymbol{\rho}\|\boldsymbol{\rho}_c)}{\partial \beta}\right|_{S(\boldsymbol{\rho}),E(\boldsymbol{\rho}),N(\boldsymbol{\rho}),\beta} &= E(\boldsymbol{\rho}) - E(\boldsymbol{\rho}_c).
\end{aligned}
\tag{6.28}
$$

How do we interpret these relations? The first relation is related to the fact that the canonical state $\boldsymbol{\rho}_c$ has maximum entropy. If the entropy $S(\boldsymbol{\rho})$ is decreasing, i.e. if the information about $\boldsymbol{\rho}$ increases, $S(\boldsymbol{\rho}\|\boldsymbol{\rho}_c)$ increases equally, as states with smaller entropy are easier to distinguish from the thermal state $\boldsymbol{\rho}_c$.

The second and the third relations can be regarded as alternative definitions for the inverse temperature β and the chemical potential μ in terms of relative entropy $S(\boldsymbol{\rho}\|\boldsymbol{\rho}_c)$ at fixed entropy $S(\boldsymbol{\rho})$. Especially the second relation is non-trivial as $E(\boldsymbol{\rho}) \neq E(\boldsymbol{\rho}_c)$ in general.

Finally, the fourth relation tells us that there exists an unique inverse temperature β, for which the two energy expectation values agree, defined through

$$\left.\frac{\partial S(\boldsymbol{\rho}\|\boldsymbol{\rho}_c)}{\partial \beta}\right|_{S(\boldsymbol{\rho}),E(\boldsymbol{\rho}),N(\boldsymbol{\rho}),\beta} = 0 \quad \Leftrightarrow \quad E(\boldsymbol{\rho}) = E(\boldsymbol{\rho}_c). \tag{6.29}$$

Through relative entropy a unique temperature can be assigned to nonequilibrium states.

We consider this inverse temperature as *optimal* in the sense that with this choice the relative entropy $S(\boldsymbol{\rho}\|\boldsymbol{\rho}_c)$ is minimized with respect to β. For any other value of β, the relative entropy is larger. In simple words, this temperature corresponds to the standard definition of temperature as if $\boldsymbol{\rho} = \boldsymbol{\rho}_c$ would hold.

GRAND CANONICAL ENSEMBLE We proceed with a similar analysis for the grand canonical ensemble describing a quantum system which is coupled to a heat and particle bath (cf. Figure 6.1c). In this case, the optimal model is given by the grand canonical density operator (we omit the construction of this state for brevity)

$$\boldsymbol{\rho}_{gc} = \frac{1}{Z_{gc}} e^{-\beta(H-\mu N)}, \tag{6.30}$$

with Z_{gc} denoting the grand canonical partition sum. Then, we find for the relative entropy between any state ρ and the optimal model ρ_{gc} the expression

$$
\begin{aligned}
S(\rho\|\rho_{\mathrm{gc}}) = & - S(\rho) + S(\rho_{\mathrm{gc}}) + \beta \left[E(\rho) - E(\rho_{\mathrm{gc}}) \right] \\
& - \beta\mu \left[N(\rho) - N(\rho_{\mathrm{gc}}) \right],
\end{aligned}
\tag{6.31}
$$

where no additional support condition has to respected. Note that strictly speaking, the two states have to correspond to the same volume V in all cases, which was not made explicit out of convenience.

Similar to the canonical case, the relative entropy decomposes into differences of extensive quantities. This also holds for the differentials

$$
\begin{aligned}
\mathrm{d}S(\rho\|\rho_{\mathrm{gc}}) = & - \mathrm{d}S(\rho) + \beta\,\mathrm{d}E(\rho) - \beta\mu\,\mathrm{d}N(\rho) \\
& + \left[E(\rho) - E(\rho_{\mathrm{gc}}) \right]\,\mathrm{d}\beta \\
& - \left[N(\rho) - N(\rho_{\mathrm{gc}}) \right]\,\mathrm{d}(\beta\mu),
\end{aligned}
\tag{6.32}
$$

leading to the (additional) partial differential relations

$$
\begin{aligned}
\left.\frac{\partial S(\rho\|\rho_{\mathrm{gc}})}{\partial N(\rho)}\right|_{S(\rho),E(\rho),\beta,\mu} &= -\beta\mu, \\
\left.\frac{\partial S(\rho\|\rho_{\mathrm{gc}})}{\partial(\beta\mu)}\right|_{S(\rho),E(\rho),N(\rho),\beta} &= -N(\rho) + N(\rho_{\mathrm{gc}}).
\end{aligned}
\tag{6.33}
$$

While the first line allows for a definition of the chemical potential μ as a partial derivative of a relative entropy, the second line shows that the optimal value for μ minimizing the relative entropy follows from choosing a model state ρ_{gc} with equal energy and particle number expectation values as the state ρ, i.e. $E(\rho) = E(\rho_{\mathrm{gc}})$ and $N(\rho) = N(\rho_{\mathrm{gc}})$, respectively.

6.2 LAWS OF THERMODYNAMICS

FIRST LAW   The first law can be regarded as energy conservation applied to thermodynamic systems and hence does not need to be formulated in terms of entropy or relative entropy. Consequently, we will omit a discussion here and focus on the two other laws of thermodynamics where entropy plays a crucial role.

SECOND LAW   We consider again the three thermodynamic ensembles (cf. Figure 6.1) and relative entropies between a state of interest ρ and the corresponding thermodynamic model state $\tilde{\rho}$ chosen such that the thermodynamic quantities, i.e. temperature T and chemical potential μ, are equal to those of the bath the system is coupled to. Then, depending on whether the system is allowed to exchange heat or

particles with the bath or not, the second law will take a specific form. In the following, we reproduce these inequalities from an information theoretic argument (see also [247]).

As pointed out in Section 3.1, the time evolution of a system (open or closed) is often described by a CPTP-map [59]. This is always the case if the global initial state is a product state, i.e. if the system and the heat bath are initially uncorrelated. However, if already initially one has classical or quantum correlations between the two systems, the time evolution map is typically only positive but not completely positive [70, 71]. The following argument holds for both types of maps and therefore explicitly allows for initial correlations. However, we are rather interested in making a general statement and hence do not specify the initial conditions.

The main assumption we use is that we consider a PTP-map $\mathcal{N}$ (possibly a CPTP-map depending on the initial conditions) which keeps the appropriate thermodynamic model state invariant, i.e. we will use $\mathcal{N}(\tilde{\rho}) = \tilde{\rho}$. This can be regarded as a definition for a stochastic time evolution. Then, one finds for the change in relative entropy during such an evolution from t_i to $t_f > t_i$

$$
\begin{aligned}
\Delta S(\rho \| \tilde{\rho}) &= S(\mathcal{N}(\rho) \| \mathcal{N}(\tilde{\rho})) - S(\rho \| \tilde{\rho}) \\
&= S(\mathcal{N}(\rho) \| \tilde{\rho}) - S(\rho \| \tilde{\rho}).
\end{aligned}
\tag{6.34}
$$

As a consequence of the monotonicity property of the quantum relative entropy (3.18), this change is always non-positive, i.e.

$$
\Delta S(\rho \| \tilde{\rho}) \leq 0,
\tag{6.35}
$$

which can be considered as a generalized second law. As usual for our relative entropy formulation, the actual state ρ does not have to be an equilibrium state itself and also we did not had to assume a quasi-stationary time evolution between equilibrium states. Rather, we ended up with the simple result that any state ρ can not deviate more from a suitable invariant thermodynamic model state $\tilde{\rho}$ over time. In fact, one can expect that in most cases the relative entropy between the two states decreases over time and converges to zero indicating that the system has thermalized with the bath.

The second law of thermodynamics is a consequence of quantum information processing.

The connection to Clausius' inequalities can be made explicit by using the expressions (6.7), (6.24) or (6.31) for the relative entropies (with the appropriate strict support conditions understood), depending on the physical situation. One can easily confirm that one finds for the change in entropy $\Delta S(\rho) = S(\mathcal{N}(\rho)) - S(\rho)$ the following relations

$$
\Delta S(\rho) \geq \begin{cases} 0 & \text{for microcanonical ensemble,} \\ \beta \Delta E(\rho) & \text{for canonical ensemble,} \\ \beta \Delta E(\rho) - \beta \mu \Delta N(\rho) & \text{for grand canonical ensemble.} \end{cases}
\tag{6.36}
$$

The microcanonical ensemble is a rather special case. The invariance condition $\mathcal{N}(\boldsymbol{\rho}_{\mathrm{m}}) = \boldsymbol{\rho}_{\mathrm{m}}$ defines the map $\mathcal{N}$ to be unital. It is well known that the change in entropy $\Delta S(\boldsymbol{\rho})$ is non-negative for unital maps [11], which is equivalent to the first line of (6.36). In contrast, the entropy might increase as well as decrease for the more general map $\mathcal{N}$ describing a system coupled to a heat or particle bath, as seen from the second and third lines of (6.36).

Let us also mention that one may derive a set of inequalities similar to (6.36) when the state $\tilde{\rho}$ is not a thermodynamic model, but instead describes a nonequilibrium steady state, such that $\mathcal{N}(\tilde{\rho}) = \tilde{\rho}$ is still fulfilled. In this case, $\Delta S(\boldsymbol{\rho})$ is also bounded from below, but the bound may not have a clear physical interpretation [262, 263].

EXAMPLE: ISOLATED THREE-STATE SYSTEM We consider an isolated three-state system with Hilbert space $\mathcal{H}$ spanned by $\{|0\rangle, |1\rangle, |2\rangle\}$ and some diagonal state ρ. The suitable model state is given by the microcanonical state $\boldsymbol{\rho}_{\mathrm{m}}$. The set of all states is shown in Figure 6.2, where $\boldsymbol{\rho}_{\mathrm{m}}$ corresponds to the central point as a consequence of the principle of minimum relative entropy.

Some generic initial state ρ can either evolve reversibly along a contour of constant entropy, which corresponds in fact to a unitary time evolution, or it might irreversibly evolve towards a state lying on a contour with higher entropy. Note that the exact position of the final states cannot be determined by the generalized second law, only the final contour is known.
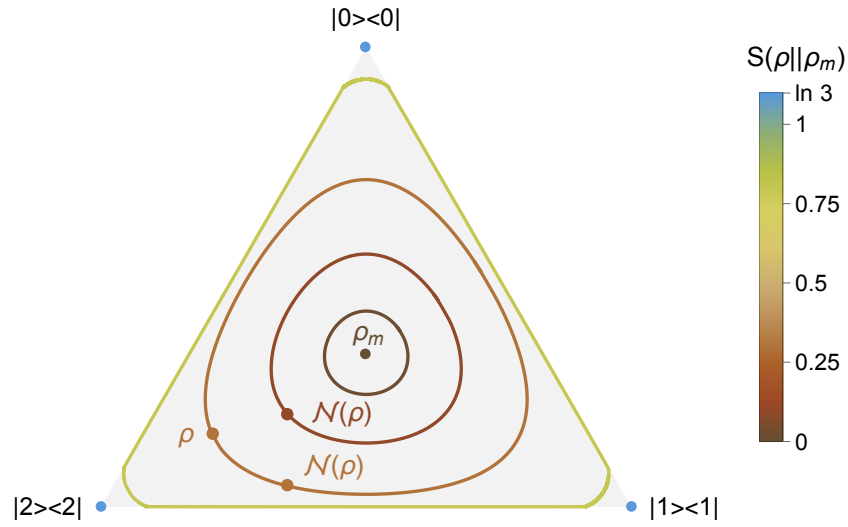

Figure 6.2: Time evolution of a state ρ describing an isolated three-state system. By the generalized second law (6.35), the relative entropy $S(\boldsymbol{\rho}\|\boldsymbol{\rho}_{\mathrm{m}})$ with respect to the microcanonical model $\boldsymbol{\rho}_{\mathrm{m}}$ can not increase under general time evolution $\mathcal{N} : \rho \to \mathcal{N}(\rho)$.

THIRD LAW   We complete our analysis with the third law of thermodynamics, which is rather closely related to the concept of entropy. Following Planck's formulation, the entropy becomes constant in the limit of vanishing temperature $T \to 0$, i.e.

$$\lim_{T \to 0} S(\tilde{\rho}) = S(\bar{\rho}), \tag{6.37}$$

for any thermodynamic model $\tilde{\rho}$. The constant is given by the entropy of the ground state $\bar{\rho}$ with respect to the Hamiltonian $\boldsymbol{H}$. Note that often the ground state is degenerate or mixed, in which case the right hand side of the latter equation evaluates to a positive number.

In terms of relative entropy, the third law can be formulated as follows: for any thermodynamic model $\tilde{\rho}$ we need to have

$$\lim_{T \to 0} S(\bar{\rho} \| \tilde{\rho}) = 0, \tag{6.38}$$

showing that the ground state $\bar{\rho}$ becomes *indistinguishable* from any thermodynamic model $\tilde{\rho}$ in the zero temperature limit $T \to 0$. Note that even if $\bar{\rho}$ is degenerate, the right hand side of (6.38) is zero.

If we choose a canonical model ρ_c, the relative entropy decomposition (6.24) shows that we have to assume $N(\rho) \equiv N(\rho_c)$ and that $E(\bar{\rho}) - E(\rho_c) \to 0$ for $T \to 0$ faster than T, such that we indeed obtain Planck's statements $S(\bar{\rho}) = S(\rho_c)$ at vanishing temperature. Similarly, we would have to assume $N(\bar{\rho}) - N(\rho_{gc}) \to 0$ for $T \to 0$ faster than T/μ for the grand canonical model ρ_{gc}.

As a concrete example, let us consider a situation where the energy eigenstates $|n\rangle$ defined through (3.25) are non-degenerate (cf. Section 3.3). Then, the canonical model state becomes (cf. (3.28))

$$\rho_c = \sum_n p_c(n) |\boldsymbol{n}\rangle \langle \boldsymbol{n}|, \quad p_c(n) = \frac{1}{Z_c} e^{-\beta E(n)}, \tag{6.39}$$

such that the unique ground state is characterized by

$$\bar{\rho} = \sum_n \bar{p}(n) |\boldsymbol{0}\rangle \langle \boldsymbol{0}|, \quad \bar{p}(n) = \delta_{0n}. \tag{6.40}$$

In this case, the relative entropy evaluates to

$$S(\bar{\rho} \| \rho_c) = S(\bar{p} \| p_c) = -\ln p_c(0). \tag{6.41}$$

Note that such a relation also holds if the model state is replaced by any state which is diagonal in the energy eigenbasis.

A short calculation shows that in the zero temperature limit $T \to 0$ we find

$$\begin{aligned}
\lim_{T \to 0} p_c(0) &= \lim_{\beta \to \infty} \frac{e^{-\beta E(n)}}{\sum_n e^{-\beta E(n)}} \\
&= 1 - \lim_{\beta \to \infty} \frac{\sum_{n>0} e^{-\beta(E(n)-E(0))}}{1 + \sum_{n>0} e^{-\beta(E(n)-E(0))}} \\
&= 1
\end{aligned} \tag{6.42}$$

as $E(n) > E(0)$ for all $n > 0$, confirming our formulation of the third law (6.38).

7 | SECOND LAW FOR RELATIVISTIC FLUIDS

The analysis is taken from [G]. S. F. proposed and supervised the project. All authors contributed equally to the development of the theory. N. D. and myself wrote large parts of earlier versions of the manuscript and all three authors finalized the draft. The figures were produced by me. As stated earlier, the article [G] is also covered in a master thesis by N. D. [248].

In Section 6.2, we have seen how a generalized second law formulated in terms of relative entropy can arise from an information theoretic argument. In this chapter, we generalize this formulation to situations where a formulation in terms of entropy is less favorable, which is in particular the case for relativistic fluids.

In general, fluid dynamics serves as an effective theory using concepts such as *local* equilibrium applicable when the evolution of a system is mainly governed by conservation laws, which are themselves consequences of continuous symmetries [264]. It turns out to be a good approximation for example as a description for the quark-gluon plasma [258, 259, 265] or for quantum magnets [266].

In the following, we specialize on *relativistic* fluid dynamics, which is built upon the symmetries of the Poincaré group. It is an interesting question how relativistic fluid dynamics arises from a local quantum field theoretic description (see e.g. [267]), which is where an information theoretic treatment could provide new insights. In particular, one may ask whether local dissipation can be understood as entanglement generation and whether such a statement can be expressed through a local second law-like inequality.

In relativistic fluid dynamics, a local second law is *postulated* in terms of an entropy current density $s^\mu(x)$ [264, 268, 269],

$$\nabla_\mu s^\mu(x) \geq 0, \tag{7.1}$$

where ∇_μ denotes the covariant derivative with respect to a general metric $g_{\mu\nu}$. The entropy current density $s^\mu(x)$ is well-defined in global equilibrium, but is rather difficult to define in nonequilibrium situations. Also, it is important to note that the entropy current $s^\mu(x)$ is a *local* quantity. Hence, if one aims at defining such a quantity starting from a local QFT, one needs to work with entanglement entropies. As discussed in Section 4.2, entanglement entropy exhibit universal UV-divergences, rendering its use less favorable.

In the affirmative, we derive second law-like inequalities in terms of relative (entanglement) entropies following from the monotonicity property (3.18). In this way, we give the local second law a solid

information theoretic foundation and avoid UV-divergences. We start with a brief introduction to relativistic fluid dynamics in Section 7.1 and discuss how a local notion of equilibrium can be defined through relative entanglement entropy in Section 7.2. Thereupon, we describe local time evolution in Section 7.3 and for such derive several second law-like inequalities in Section 7.4.

7.1 RELATIVISTIC FLUID DYNAMICS

CONSERVATION LAWS   We formulate relativistic fluid dynamics in general coordinates x^μ, such that we have to work with a general metric $g_{\mu\nu}(x)$ and covariant derivatives ∇_μ. Then, conservation laws for energy and momentum are in fact a consequence of diffeomorphism invariance and can be subsumed as

$$\nabla_\mu T^{\mu\nu}(x) = 0, \tag{7.2}$$

where $T^{\mu\nu}(x)$ denotes the classical energy-momentum tensor.

Additionally, we assume that the theory exhibits a $U(1)$ symmetry, leading to the conservation of a particle number current

$$\nabla_\mu N^\mu(x) = 0. \tag{7.3}$$

Apparently, (7.2) and (7.3) provide $d + 1$ equations, while $T^{\mu\nu}(x)$ has $d(d+1)/2$ and $N^\mu(x)$ has d components. The fundamental assumption of relativistic fluid dynamics is now to parameterize these two tensors in terms of $d + 1$ variables, which may chosen to be the temperature current

Relativistic fluid dynamics emerges from symmetry considerations.

$$\beta^\mu(x) = \frac{u^\mu(x)}{T(x)}, \tag{7.4}$$

where $u^\mu(x)$ denotes the local fluid velocity and $T(x)$ is a local temperature field, and the scalar field

$$\alpha(x) = \frac{\mu(x)}{T(x)}, \tag{7.5}$$

with $\mu(x)$ being the local chemical potential field. The two variables $\beta^\mu(x)$ and $\alpha(x)$ are then determined by the two conservation laws (7.2) and (7.3).

KILLING EQUATIONS IN EQUILIBRIUM   The appearance of these fields can be further motivated from relativistic thermodynamics. To that end, we consider an entropy current $s^\mu(x)$, which is expected to fulfill a local second law in the form of

$$\nabla_\mu s^\mu(x) \geq 0, \tag{7.6}$$

with equality in thermal equilibrium. Note again that this is a phenomenological assumption, which we wish to understand better. Especially, it is not clear whether an entropy current $s^\mu(x)$ can be defined in general nonequilibrium situations at all or how it is linked to local entropies in a QFT.

However, in global thermal equilibrium, we can write the entropy current as a function of the other two conserved quantities $s^\mu = s^\mu(T^{\nu\lambda}, N^\sigma)$, such that

$$\nabla_\mu s^\mu = \frac{\partial s^\mu}{\partial T^{\nu\lambda}} \nabla_\mu T^{\nu\lambda} + \frac{\partial s^\mu}{\partial N^\nu} \nabla_\mu N^\nu. \tag{7.7}$$

Then, combining (7.2) and (7.3) with (7.6) at equality leads to the covariant generalizations of the usual thermodynamic relations

$$\frac{\partial s^\mu}{\partial T^{\nu\lambda}} = -\beta_\lambda \delta^\mu_\nu, \quad \frac{\partial s^\mu}{\partial N^\nu} = -\alpha \delta^\mu_\nu. \tag{7.8}$$

In thermal equilibrium, the two fields $\beta^\mu(x)$ and $\alpha(x)$ can be constrained from the requirement that $\nabla_\mu s^\mu$ must be stationary, i.e.

$$\nabla_\mu ds^\mu = -\nabla_\mu \beta_\nu dT^{\mu\nu} + -\partial_\mu \alpha dN^\mu = 0, \tag{7.9}$$

implying that $\beta^\mu(x)$ has to be a Killing vector field and that $\alpha(x)$ is constant,

$$\nabla_\mu \beta_\nu + \nabla_\nu \beta_\mu = 0, \quad \partial_\mu \alpha = 0. \tag{7.10}$$

In fact, one may define equilibrium by the latter two conditions. Note that for a Minkowskian metric $g_{\mu\nu} = \eta_{\mu\nu}$, the Killing vector field condition translates into the temperature being constant.

In a generic nonequilibrium situation, the two variables $\beta^\mu(x)$ and $\alpha(x)$ are not defined in a unique way. For example, one may choose the fluid velocity $u^\mu(x)$ to be parallel to the energy flow (the so-called Landau frame definition) or parallel to the particle number flow (the Eckart frame definition). Later, we will be concerned with the notion of *local* equilibrium, where (7.10) will not hold anymore.

GRADIENT EXPANSION   Let us briefly discuss the gradient expansion, which is typically used within a phenomenological approach to fluid dynamics to account for local deviations from equilibrium. Most importantly, fluid dynamics uses the concept of equilibrium locally as an approximation. In the hydrodynamic gradient expansion (see [264]), i.e. an expansion in terms of gradients of fluid velocity $u^\mu(x)$, temperature $\beta^\mu(x)$, etc., this approximation becomes exact when these gradients vanish, in which case we recover global equilibrium defined through (7.10). Depending on how many terms in this derivative expansion are kept, the energy-momentum tensor $T^{\mu\nu}(x)$ and the particle number current $N^\mu(x)$ take specific forms. For example, in lowest order, one recovers ideal fluid dynamics, which will be discussed in more detail

in Section 7.2. Higher-order terms can be seen as corrections to the ideal fluid approximation and allow to describe dissipative effects.

Also, one may describe dissipative effects by introducing additional parameters, which vanish in global equilibrium, but give corrections to this approximation otherwise. For example, one may consider shear stress $\pi^\nu(x)$, bulk viscous pressure $\pi_{\text{bulk}}(x)$ or the diffusion current $\nu^\mu(x)$ within the theory of Israel and Stewart [268], which then allow to parameterize the energy-momentum tensor $T^{\mu\nu}(x)$ and the particle number current $N^\nu(x)$ together with the aforementioned fields in terms of the so-called constitutive relations. Although the number of parameters is larger compared to the equilibrium case, it is still small compared to the degrees of freedom of a general nonequilibrium state, which emphasizing the approximation character of a local fluid-like description.

7.2 GLOBAL AND LOCAL EQUILIBRIUM

GLOBAL EQUILIBRIUM Instead of pursuing the phenomenological approach, we propose to use relative entanglement entropy to justify a local equilibrium approximation. To that end, we ask how the local parameters are related to local expectation values of the respective quantum operators for given local quantum states and which role entanglement plays for a fluid-like description to work. We begin with a brief review of global equilibrium quantum states.

Let us consider a foliation of spacetime in terms of a family of Cauchy hypersurfaces $\Sigma(\tau)$ equipped with a unit time-like normal vector $n^\mu(x)$ fulfilling $n_\mu n^\mu = -1$. On every hypersurface $\Sigma(\tau)$, we may define the global equilibrium state as the covariant generalization of (6.30) with the conditions (7.10) understood, which gives

$$\tilde{\rho} = \frac{1}{Z} \exp\left[-\int_{\Sigma(\tau)} \mathrm{d}\Sigma_\mu \left(\beta_\nu T^{\mu\nu} + \alpha N^\mu \right) \right] \tag{7.11}$$

where Z is the partition function

$$Z = \mathrm{Tr}\left\{ \exp\left[-\int_{\Sigma(\tau)} \mathrm{d}\Sigma_\mu \left(\beta_\nu T^{\mu\nu} + \alpha N^\mu \right) \right] \right\}. \tag{7.12}$$

Therein, the hypersurface area element $\mathrm{d}\Sigma_\mu$ is given by

$$\mathrm{d}\Sigma_\mu = \mathrm{d}^{d-1}y \sqrt{|\det h_{\mu\nu}|}\, n_\mu, \tag{7.13}$$

with n^μ being future oriented such that $n^0 > 0$ (similar to a fluid velocity) and where $h_{\mu\nu}$ denotes the induced metric on the hypersurface $\Sigma(\tau)$. Note also that the global equilibrium state (7.11) contains the *quantum* operators $T^{\mu\nu}(x)$ and $N^\mu(x)$.

As discussed above, such a state corresponds to the zeroth order in the gradient expansion. In this case, there exists a unique fluid frame

(i.e. the Landau and Eckart frames are equivalent) in which (7.4) and (7.5) become exact. For the corresponding expectation values of the energy-momentum tensor and the particle number current one finds the constitutive relations [264]

$$T^{\mu\nu} = \epsilon u^\mu u^\nu + p\Delta^{\mu\nu}, \quad N^\mu = nu^\mu, \quad s^\mu = su^\mu, \tag{7.14}$$

with s being the entropy density, ϵ the energy density, n the particle number density, p the pressure and $\Delta^{\mu\nu} = g^{\mu\nu} + u^\mu u^\nu$ a projector orthogonal to u^μ.

FLUID CELL APPROXIMATION WITH RELATIVE ENTROPY   Let us now split the global hypersurface Σ into a finite subregion 1 and its complement 2 such that $\Sigma = 1 \cup 2 \equiv 12$ and $1 \cap 2 = \varnothing$, for example a ball-shaped region of radius R around $\vec{x}_1 \in \Sigma$ at some time x_1^0. We consider a generic global quantum state ρ_{12}. This state is assumed to be out-of-global equilibrium in general. To the local region 1 we associate a local state

$$\rho_1 = \mathrm{Tr}_2\{\rho_{12}\}, \tag{7.15}$$

Then, the expectation value of a local observable O_1 can be computed from

$$O_1 = \mathrm{Tr}_1\{\rho_1 O_1\}. \tag{7.16}$$

As there exist many different global states ρ_{12}, which have the same local states ρ_1 in the sense of (7.15), one can compute local expectation values also when assuming the two regions 1 and 2 to be uncorrelated, i.e. $\rho_{12} = \rho_1 \otimes \rho_2 = \mathrm{Tr}_2\{\rho_{12}\} \otimes \mathrm{Tr}_1\{\rho_{12}\}$. It may be possible to understand the local equilibrium approximation in this way. More precisely, a fluid-like description in terms of local parameters may work rather well when non-local effects such as entanglement between the two regions 1 and 2 are neglected to some extent.

Note that such a statement is rather informal and should be investigated with much more mathematical rigor in the future. However, it is not needed for what follows. Rather, it is essential that local parameters can be compatible with many global states, giving us some freedom to choose a specific one. In particular, we may consider the global equilibrium state $\tilde{\rho}_{12}$ defined in (7.11), specified with parameters $\beta^\mu(x)$ and $\alpha(x)$ which fulfill (7.10), and a corresponding local state

$$\tilde{\rho}_1 = \mathrm{Tr}_2\{\tilde{\rho}_{12}\} \tag{7.17}$$

in region 1. Based on this, we can define local equilibrium through an argument involving distinguishability. We say that the generic state ρ_{12} is in local equilibrium in the region 1 around $\vec{x}_1$, such that we

may associate local parameters $\beta^\mu(x)$ and $\alpha(x)$, when the relative entanglement entropy between ρ_1 and $\tilde{\rho}_1$ becomes small

$$S(\rho_1\|\tilde{\rho}_1) \to 0, \tag{7.18}$$

such that ρ_1 and $\tilde{\rho}_1$ can be considered to be locally *indistinguishable*. Hence, we may quantify the validity of a local equilibrium approximation from a quantum information theoretic perspective in terms of a relative entanglement entropy.

Also, $S(\rho_1\|\tilde{\rho}_1)$ being small allows us to define what we call the *fluid cell approximation*. In fact, (7.18) indicates the existence of an intermediate scale, where the size R of the region 1 is large enough compared to an UV regulator scale and at the same time small enough compared to the scale where long-range effects become relevant. More precisely, for a global nonequilibrium state ρ_{12} we have $S(\rho_1\|\tilde{\rho}_1) \gg 0$ for sufficiently large R, but $S(\rho_1\|\tilde{\rho}_1)$ becomes small (i.e. receives a non-trivial upper bound) on the size of a fluid cell.

LOCAL EQUILIBRIUM The fluid cell approximation may be extended from one local fluid cell to the entire hypersurface Σ by assigning local parameters $\beta^\mu(x)$ and $\alpha(x)$ to every fluid cell region where (7.18) is approximately fulfilled when adjusting the model $\tilde{\rho}$ accordingly. One might choose a global state of the form (7.11), but without the conditions (7.10) on $\beta^\mu(x)$ and $\alpha(x)$, i.e.

A local equilibrium state may be constructed from local fluid cells.

$$\tilde{\rho}_{12} = \frac{1}{Z} \exp\left[-\int_{\Sigma(\tau)} d\Sigma_\mu \left(\beta_\nu T^{\mu\nu} + \alpha N^\mu\right)\right] \tag{7.19}$$

with

$$Z = \mathrm{Tr}\left\{\exp\left[-\int_{\Sigma(\tau)} d\Sigma_\mu \left(\beta_\nu T^{\mu\nu} + \alpha N^\mu\right)\right]\right\}. \tag{7.20}$$

We may call (7.19) a *local equilibrium state*, in the sense that (7.18) is fulfilled for every local fluid cell. Note that this does not imply $\rho_{12} = \tilde{\rho}_{12}$ on a global level, but rather represents some coarse-graining of ρ_{12} on the scale set by a fluid cell. In particular, ρ_{12} may be pure, while $\tilde{\rho}_{12}$ is mixed and they can still be hard to distinguish locally.

As a side note, one may obtain a density operator of the form (7.19) following Zubarev's approach, using the maximum entropy principle on every hypersurface with given constraints on the expectation values of the energy-momentum tensor and the particle number [270, 271].

7.3 LOCAL TIME EVOLUTION IN DOUBLE LIGHT CONE REGION

LOCAL TIME EVOLUTION Ultimately, we want to understand *local* time evolution of an open quantum system on a scale where a fluid-like description is applicable. To that end, one could investigate the

local system directly in terms of local density operators, which requires detailed considerations taking care of the subtleties in the field theory limit. We leave this task for future work [272, 273].

In the following, we propose an alternative strategy to overcome the conceptual difficulties of local states. In particular, we make statements about a general global state ρ_{12} by comparing it to the global or local equilibrium states $\tilde{\rho}_{12}$ discussed above, with the caveat that the global time evolution is chosen such that it is *effectively local*.

To explicitly focus on time evolution only in a local region 1, we choose a family of hypersurfaces $\Sigma(\tau)$ which differ only in region $1(\tau)$, but remain stationary in the complement region $2 \neq 2(\tau)$. Then, the time evolution is indeed essentially local. For such a construction to work, we need to have the normal vectors $n_\mu(x)$ on each surface $\Sigma(\tau)$ pointing into a timelike (or lightlike as a limit) direction. Moreover, they have to be ordered in a way compatible with time running forward only. Apart from these two constraints, the family $\Sigma(\tau)$ can be chosen rather freely.

With this idea in mind, we ask how the phenomenological local second law (7.6) can be understood from a quantum information theoretic point of view. In a local spacetime region Ω, which can be chosen such that its boundaries are given by local regions on two subsequent hypersurfaces $\Sigma(\tau_f)$ and $\Sigma(\tau_i)$, i.e. $\partial\Omega = 1(\tau_f) \cup 1(\tau_i)$, the integrated form of (7.6) reads

$$\int_\Omega \mathrm{d}^d x \sqrt{\det g_{\mu\nu}} \nabla_\mu s^\mu(x) = \oint_{\partial\Omega} \mathrm{d}\Sigma_\mu s^\mu(x) \geq 0. \tag{7.21}$$

Therein, we have used again the hypersurface area element $\mathrm{d}\Sigma_\mu$ as given in (7.13). However, it is the first time we have to deal with a closed surface integral, where some subtleties regarding the orientation of the normal vector have to be specified. For the chosen metric signature $(-,+,+,+)$, $n^\mu(x)$ does unambiguously point to the outside if it is spacelike. If it is timelike instead, the normal vector is pointing to the inside (see e.g. [274] Sec. B.2, [275] Sec. 16 or [276] for more details). For the closed surface $\partial\Omega$, for which $n^\mu(x)$ points in a time-like direction, the orientation of $n^\mu(x)$ flips between two subsequent hypersurfaces. More precisely, $n^0 > 0$ for the past and $n^0 < 0$ for the future hypersurfaces, respectively.

Interestingly, the generalized divergence theorem implies that (7.21) follows from (7.6) for any spacetime region Ω with properly chosen boundaries $\partial\Omega$. At the same time, (7.21) can also imply (7.6) if it can be proven for an arbitrarily small region around a point $x \in \Omega$. Nevertheless, it remains unclear how (or if at all) the left hand side of (7.21) is related to a von Neumann entropy of a nonequilibrium density operator. To overcome this problem, we work with relative entropies in Section 7.4, such that we never have to define an entropy current outside of equilibrium.

DOUBLE LIGHT CONE A geometry which represents the aforementioned situation is the double light cone as sketched in Figure 7.1 for $d = 1 + 1$ spacetime dimensions. The double light cone is constructed from the future light cone of some point p in the past and the past light cone of another point q in the future. In this setup, the light cone associated with p can be understood as an initial hypersurface, while the light cone associated with q as a final hypersurface with respect to the time evolution inside the light cone. Hence, the dynamics are restricted to the region within the double light cone and the non-local information from outside the double light cone is entirely encoded in the initial hypersurface. In this sense, there is no transfer of quantum information through the boundaries. As a consequence, entropy production, possible caused by an interaction with another system, will be localized within the double light cone.
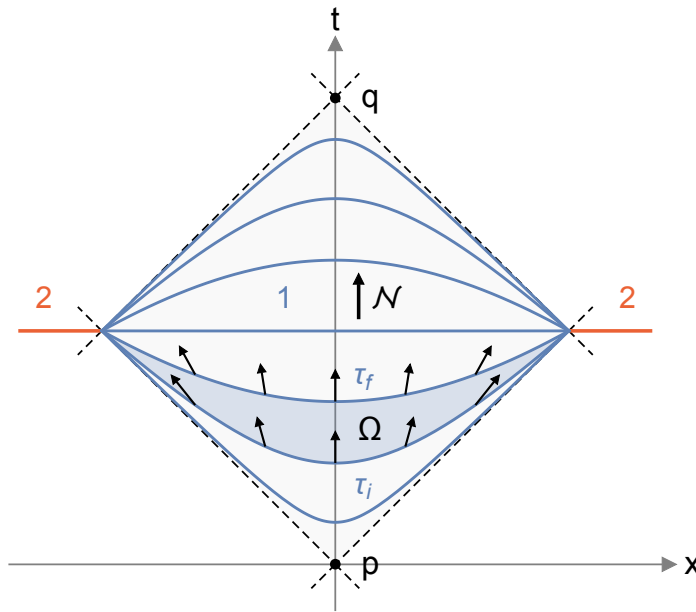
The double light cone is causally closed

Figure 7.1: A double light cone bounded by the future light cone of p and the past light cone of q (dashed black lines). Local time evolution is modeled by a family of Cauchy hypersurfaces $\Sigma(\tau)$ which differ only in region $1(\tau)$ (solid blue curves) within the double light cone, while the complement 2 (solid red lines) remains static. Mathematically, the time evolution is represented by the map $\mathcal{N}$, which is specified depending on whether the system of interest is isolated or coupled to another system. Generally, two subsequent hypersurfaces $\Sigma(\tau_f)$ and $\Sigma(\tau_i)$ enclose a spacetime region Ω (shaded blue region), where each hypersurface is equipped with timelike normal vectors $n^\mu(x)$ (indicated by small black arrows).

As shown in Figure 7.1, one can choose a foliation of spacetime $\Sigma(\tau)$ with timelike unit normal $n^\mu(x)$, where τ plays the role of a generalized time coordinate, such that the dynamics are constrained to a local region $1(\tau) \subseteq \Sigma(\tau)$ within the double light cone. The

complement region $2 \subseteq \Sigma(\tau)$ remains static. The two regions intersect where the two light cones intersect, which may be considered as fixed points as these points are part of all hypersurfaces $\Sigma(\tau)$. For $d = 1 + 1$ spacetime dimensions, the fixed points are actually points, while for $d = 3 + 1$ the fixed points are given by 2-spheres.

COUPLING TO A BATH FLUID   The time evolution of an arbitrary global state ρ_{12} from one hypersurface $\Sigma(\tau_i)$ to the next $\Sigma(\tau_f)$ is generated by a map $\mathcal{N}$ via

$$\mathcal{N} : \rho_{12}(\tau_i) \rightarrow \mathcal{N}(\rho_{12}(\tau_i)) = \rho_{12}(\tau_f). \tag{7.22}$$

If the system under consideration is closed, $\mathcal{N}$ is unitary. As this corresponds to a rather uninteresting situation, we couple the field of interest locally to a so-called bath fluid through an interaction parameter λ, such that we end up with an open system whose time evolution is again non-trivial. Hence, despite the fact that we are interested in a causally closed region, quantum information is not conserved, but instead constrained by a local second law-like inequality.

Then, the map $\mathcal{N}$ is a CPTP- or at least a PTP-map and we choose it such that it keeps a particular equilibrium reference state $\tilde{\rho}_{12}$ form invariant similar to the considerations in Section 6.2. Also, in the limit $\lambda \rightarrow 0$, this map becomes unitary again. In the following, we derive second law-like inequalities for an arbitrary global quantum state ρ_{12} for three different physical situations encoded in the choice of the invariant reference state $\tilde{\rho}_{12}$.

7.4   SECOND LAW–LIKE INEQUALITIES

BATH FLUID IN GLOBAL EQUILIBRIUM   We start with the simplest case, i.e. a bath fluid in global equilibrium, such that a suitable model state is given by (7.11) with (7.10) understood. An example for such a situation is a gas consisting of electron and positrons which is coupled to a bath of electromagnetic radiation at some temperature $T(x)$ and vanishing chemical potential $\mu(x) = 0$. In this case, the coupling λ is given by the elementary charge e.

For a global equilibrium state $\tilde{\rho}_{12}$, the corresponding von Neumann entropy $S(\tilde{\rho}_{12})$ decomposes according to

$$S(\tilde{\rho}_{12}) = \ln Z + \int d\Sigma_\mu \left(\beta_\nu T^{\mu\nu}(\tilde{\rho}_{12}) + \alpha N^\mu(\tilde{\rho}_{12}) \right) \tag{7.23}$$

where we have introduced the expectation values

$$T^{\mu\nu}(\tilde{\rho}_{12}) = \text{Tr}\left\{ \tilde{\rho}_{12} T^{\mu\nu} \right\}, \quad N^\mu(\tilde{\rho}_{12}) = \text{Tr}\left\{ \tilde{\rho}_{12} N^\mu \right\}, \tag{7.24}$$

which are assumed to be renormalized such that they vanish in the vacuum $T = \mu = 0$.

In global equilibrium, $S(\tilde{\rho}_{12})$ is extensive and can be written as an integral over a local entropy current $s^\mu(x)$, such that

$$S(\tilde{\rho}_{12}) = -\int d\Sigma_\mu\, s^\mu(\tilde{\rho}_{12}), \tag{7.25}$$

where the minus sign in the last line is a consequence of $d\Sigma_\mu$ and $s^\mu(x)$ being future oriented with $d\Sigma_0, s^0(x) > 0$ together with the metric signature convention $(-,+,+,+)$.

Also the Schwinger functional (at vanishing sources) can be written as an integral

$$W = \ln Z = -\int d\Sigma_\mu\, p\beta^\mu, \tag{7.26}$$

where p denotes the pressure with $p = 0$, i.e. $Z = 1$, in the vacuum $T = \mu = 0$. Hence, the entropy current density reads

$$s^\mu = -\beta_\nu T^{\mu\nu} - \alpha N^\mu + p\beta^\mu. \tag{7.27}$$

This expression is consistent with the constitutive relation $s^\mu = su^\mu$ for an ideal fluid (7.14), which follows from the Gibbs-Durham relation $\epsilon + p = sT + \mu n$ together with the other two relations in (7.14).

Next, we consider the quantum relative entropy between an arbitrary global state ρ_{12} and the global equilibrium state $\tilde{\rho}_{12}$. We find a decomposition similar to (6.31), i.e.

$$\begin{aligned}
&S(\rho_{12}\|\tilde{\rho}_{12})\\
&= -S(\rho_{12}) + \ln Z + \mathrm{Tr}\left\{\rho_{12}\int d\Sigma_\mu\left(\beta_\nu T^{\mu\nu} + \alpha N^\mu\right)\right\}\\
&= -S(\rho_{12}) + \int d\Sigma_\mu\left(-s^\mu(\tilde{\rho}_{12})\right.\\
&\quad\left. + \beta_\nu\left[T^{\mu\nu}(\rho_{12}) - T^{\mu\nu}(\tilde{\rho}_{12})\right] + \alpha\left[N^\mu(\rho_{12}) - N^\mu(\tilde{\rho}_{12})\right]\right),
\end{aligned} \tag{7.28}$$

where, except for the non-local quantity $S(\rho_{12})$, all quantities can be written as local integrals over a full Cauchy hypersurface $\Sigma(\tau)$.

Between two times τ_i and τ_f, the relative entropy changes by an amount of

$$\Delta S(\rho_{12}\|\tilde{\rho}_{12}) = S(\rho_{12}(\tau_f)\|\tilde{\rho}_{12}(\tau_f)) - S(\rho_{12}(\tau_i)\|\tilde{\rho}_{12}(\tau_i)). \tag{7.29}$$

Through the coupling to the bath fluid, the time evolution is given by a PTP-map (or possibly CPTP map) $\mathcal{N}$ which leaves $\tilde{\rho}_{12}(\tau_i)$ invariant in the sense that $\tilde{\rho}_{12}(\tau_f)$ is of the same form, but defined on a different hypersurface $\Sigma(\tau_f)$. By the monotonicity of the quantum relative entropy (3.18), we can conclude that

$$\Delta S(\rho_{12}\|\tilde{\rho}_{12}) \leq 0 \tag{7.30}$$

with equality for a unitary time evolution, which corresponds to vanishing coupling $\lambda = 0$. This inequality can be considered as a global second law for relativistic fluids.

To derive a local form of the latter inequality, we use (7.28). Then,

$$
\begin{aligned}
\Delta S(\boldsymbol{\rho}_{12}\|\tilde{\boldsymbol{\rho}}_{12}) = {} & -\Delta S(\boldsymbol{\rho}_{12}) - \oint_{\partial\Omega} \mathrm{d}\Sigma_\mu \Big(-s^\mu(\tilde{\boldsymbol{\rho}}_{12}) \\
& + \beta_\nu \left[T^{\mu\nu}(\boldsymbol{\rho}_{12}) - T^{\mu\nu}(\tilde{\boldsymbol{\rho}}_{12}) \right] \\
& + \alpha \left[N^\mu(\boldsymbol{\rho}_{12}) - N^\mu(\tilde{\boldsymbol{\rho}}_{12}) \right] \Big),
\end{aligned}
\tag{7.31}
$$

where the integral goes over the closed surface $\partial\Omega = 1(\tau_{\mathrm{f}}) \cup 1(\tau_{\mathrm{i}})$ and where we introduced the change in entropy as

$$
\Delta S(\boldsymbol{\rho}_{12}) = S(\boldsymbol{\rho}_{12}(\tau_{\mathrm{f}})) - S(\boldsymbol{\rho}_{12}(\tau_{\mathrm{i}})).
\tag{7.32}
$$

Note that by convention, we have an additional minus sign in front of the integral (cf. discussion around (7.21)). Note also that all contributions from region 2 canceled as this region remains static.

Although the von Neumann entropy $S(\boldsymbol{\rho}_{12})$ is non-local in general, the change in entropy $\Delta S(\boldsymbol{\rho}_{12})$ *is* local since it is caused by the interaction with the bath fluid, which is assumed to be local in a relativistic QFT. Hence, we may write

$$
\Delta S(\boldsymbol{\rho}_{12}) = \int_\Omega \mathrm{d}^d x \sqrt{\det g_{\mu\nu}}\, \mathfrak{s}(\boldsymbol{\rho}_{12})(x),
\tag{7.33}
$$

where the local quantity $\mathfrak{s}(\boldsymbol{\rho}_{12})$ is well-defined for *all* density operators and does not rely on assuming equilibrium as it was the case for the entropy current $s^\mu(x)$.

Then, (7.31) becomes with the generalized divergence theorem

$$
\begin{aligned}
0 \geq {} & \Delta S(\boldsymbol{\rho}_{12}\|\tilde{\boldsymbol{\rho}}_{12}) \\
= {} & \int_\Omega \mathrm{d}^d x \sqrt{\det g_{\mu\nu}} \Big(-\mathfrak{s}(\boldsymbol{\rho}_{12}) - \beta_\nu \nabla_\mu T^{\mu\nu}(\boldsymbol{\rho}_{12}) \\
& \qquad\qquad - \alpha \nabla_\mu N^\mu(\boldsymbol{\rho}_{12}) \Big),
\end{aligned}
\tag{7.34}
$$

where we used the equilibrium conditions (7.10) together with the conservation laws (7.2), (7.3) and (7.6) for the global equilibrium state $\tilde{\boldsymbol{\rho}}_{12}$. Since the latter inequality must be fulfilled for any choice of the Cauchy hypersurfaces $\Sigma(\tau)$, its local form must be fulfilled as well.

The local change in entropy is well-defined outside of equilibrium and constrained by a second law-like inequality.

Hence, we obtain the local second law-like inequality

$$
\mathfrak{s}(\boldsymbol{\rho}_{12}) \geq -\beta_\nu \nabla_\mu T^{\mu\nu}(\boldsymbol{\rho}_{12}) - \alpha \nabla_\mu N^\mu(\boldsymbol{\rho}_{12}),
\tag{7.35}
$$

which corresponds to the local form of the third line in (6.36).

Let us also formulate a local form of (7.30). Repeating the argument that interactions are local we can write

$$
0 \geq \Delta S(\boldsymbol{\rho}_{12}\|\tilde{\boldsymbol{\rho}}_{12}) = \int_\Omega \mathrm{d}^d x \sqrt{\det g_{\mu\nu}}\, \mathfrak{s}(\boldsymbol{\rho}_{12}\|\tilde{\boldsymbol{\rho}}_{12}),
\tag{7.36}
$$

such that we obtain the local statement

$$
\mathfrak{s}(\boldsymbol{\rho}_{12}\|\tilde{\boldsymbol{\rho}}_{12}) \leq 0,
\tag{7.37}
$$

showing that the local relative entropy production is non-positive. In this sense, states become locally less distinguishable over time.

RELATIVE ENTANGLEMENT ENTROPY FORMULATION   Having found a local form of the second law in terms of a local change in entropy (7.35) and also in relative entropy (7.37), let us specify further to local density operators in the region 1 defined through (7.15). The analog of (7.30) in terms of relative entanglement entropy then reads

$$\Delta S(\boldsymbol{\rho}_1 \| \tilde{\boldsymbol{\rho}}_1) = S(\boldsymbol{\rho}_1(\tau_f) \| \tilde{\boldsymbol{\rho}}_1(\tau_f)) - S(\boldsymbol{\rho}_1(\tau_i) \| \tilde{\boldsymbol{\rho}}_1(\tau_i)) \leq 0, \qquad (7.38)$$

where we used that the local states evolve in time according to modified operators $\mathcal{N}_1$ such that

$$\begin{aligned}
\boldsymbol{\rho}_1(\tau_f) &= \mathcal{N}_1(\boldsymbol{\rho}_1(\tau_i)) \\
&= \mathcal{N}_1(\mathrm{Tr}_2\{\boldsymbol{\rho}(\tau_i)\}) \\
&= \mathrm{Tr}_2\{\mathcal{N}(\boldsymbol{\rho}(\tau_i))\} \\
&= \mathrm{Tr}_2\{\boldsymbol{\rho}(\tau_f)\},
\end{aligned} \qquad (7.39)$$

showing that $\mathcal{N}_1$ is also a CPTP- or at least a PTP-map, allowing us to apply the monotonicity property (3.18).

For local interactions with the bath fluid we can write

$$0 \geq \Delta S(\boldsymbol{\rho}_1 \| \tilde{\boldsymbol{\rho}}_1) = \int_\Omega \mathrm{d}^d x \sqrt{\det g_{\mu\nu}} \, \mathfrak{s}(\boldsymbol{\rho}_1 \| \tilde{\boldsymbol{\rho}}_1). \qquad (7.40)$$

Finally, let us remark that by construction it is highly plausible that the local changes in relative entropy and relative entanglement entropy agree,

$$\mathfrak{s}(\boldsymbol{\rho}_1 \| \tilde{\boldsymbol{\rho}}_1) = \mathfrak{s}(\boldsymbol{\rho}_{12} \| \tilde{\boldsymbol{\rho}}_{12}), \qquad (7.41)$$

such that also

$$\Delta S(\boldsymbol{\rho}_1 \| \tilde{\boldsymbol{\rho}}_1) = \Delta S(\boldsymbol{\rho}_{12} \| \tilde{\boldsymbol{\rho}}_{12}). \qquad (7.42)$$

This would allow to formulate a local second law-like inequality in terms of local states. However, a more detailed investigation including a formal proof of the latter two equalities is required. We leave such an analysis for future work.

BATH FLUID IN LOCAL EQUILIBRIUM   We now generalize the latter considerations to the case where the bath fluid is of local equilibrium form (7.19) without the conditions (7.10). Again, we assume the time evolution map $\mathcal{N}$ to be a CPTP- or at least PTP-map which leaves the form of the global reference state $\tilde{\boldsymbol{\rho}}_{12}$ invariant.

In this scenario, the von Neumann entropy $S(\tilde{\boldsymbol{\rho}}_{12})$ is still of the form (7.23), i.e. it can be written in terms of a Schwinger functional $W = \ln Z$ and expectation values of energy-momentum tensor $T^{\mu\nu}(\tilde{\boldsymbol{\rho}}_{12})$ and particle number current $N^\mu(\tilde{\boldsymbol{\rho}}_{12})$. In [277] it was argued that W is extensive also in this case, such that it can be written as

$$W = \ln Z = - \int \mathrm{d}\Sigma_\mu \, w^\mu(\tilde{\boldsymbol{\rho}}_{12}), \qquad (7.43)$$

allowing us to define a nonequilibrium entropy current

$$s^\mu(\tilde{\boldsymbol{\rho}}_{12}) = -\beta_\nu T^{\mu\nu}(\tilde{\boldsymbol{\rho}}_{12}) - \alpha N^\mu(\tilde{\boldsymbol{\rho}}_{12}) + w^\mu(\tilde{\boldsymbol{\rho}}_{12}), \tag{7.44}$$

such that (7.25) holds. Note that for $w^\mu = p\beta^\mu$ we indeed obtain from the latter equation the entropy current in the Landau frame within the first order gradient expansion, which may be seen as a justification. However, a more detailed investigation from a field theoretic perspective may be in order.

Let us now consider the change in relative entropy $\Delta S(\boldsymbol{\rho}_{12}\|\tilde{\boldsymbol{\rho}}_{12})$ again. As before, the relative entropy admits a decomposition of the form (7.31), where the two parameter fields $\beta^\mu(x)$ and $\alpha(x)$ do not fulfill the global equilibrium conditions (7.10) anymore. Then, by using monotonicity of quantum relative entropy (3.18) together with the locality of the change in entropy (7.32) we obtain

$$\begin{aligned}
0 \geq{}& S(\boldsymbol{\rho}_{12}\|\tilde{\boldsymbol{\rho}}_{12}) \\
={}& \int_\Omega \mathrm{d}^d x \sqrt{\det g_{\mu\nu}} \Big(-\mathfrak{s}(\boldsymbol{\rho}_{12}) + \nabla_\mu s^\mu(\tilde{\boldsymbol{\rho}}_{12}) \\
&- \beta_\nu \nabla_\mu \left[T^{\mu\nu}(\boldsymbol{\rho}_{12}) - T^{\mu\nu}(\tilde{\boldsymbol{\rho}}_{12}) \right] - \alpha \nabla_\mu \left[N^\mu(\boldsymbol{\rho}_{12}) - N^\mu(\tilde{\boldsymbol{\rho}}_{12}) \right] \\
&- (\nabla_\mu \beta_\nu) \left[T^{\mu\nu}(\boldsymbol{\rho}_{12}) - T^{\mu\nu}(\tilde{\boldsymbol{\rho}}_{12}) \right] - (\partial_\mu \alpha) \left[N^\mu(\boldsymbol{\rho}_{12}) - N^\mu(\tilde{\boldsymbol{\rho}}_{12}) \right] \Big).
\end{aligned} \tag{7.45}$$

By using the definition of the entropy current (7.44), this simplifies to

$$\begin{aligned}
0 \geq{}& S(\boldsymbol{\rho}_{12}\|\tilde{\boldsymbol{\rho}}_{12}) \\
={}& \int_\Omega \mathrm{d}^d x \sqrt{\det g_{\mu\nu}} \Big(-\mathfrak{s}(\boldsymbol{\rho}_{12}) + \nabla_\mu w^\mu(\tilde{\boldsymbol{\rho}}_{12}) \\
&- \beta_\nu \nabla_\mu T^{\mu\nu}(\boldsymbol{\rho}_{12}) - \alpha \nabla_\mu N^\mu(\boldsymbol{\rho}_{12}) \\
&- (\nabla_\mu \beta_\nu) T^{\mu\nu}(\boldsymbol{\rho}_{12}) - (\partial_\mu \alpha) N^\mu(\boldsymbol{\rho}_{12}) \Big).
\end{aligned} \tag{7.46}$$

If no energy, momentum or particles are exchanged between the field of interest and the bath fluid, the third lines in the latter two equations drop out, such that we may write a local change in relative entropy as

$$\begin{aligned}
0 \geq{}& \mathfrak{s}(\boldsymbol{\rho}_{12}\|\tilde{\boldsymbol{\rho}}_{12}) \\
={}& -\mathfrak{s}(\boldsymbol{\rho}_{12}) + \nabla_\mu s^\mu(\tilde{\boldsymbol{\rho}}_{12}) \\
&- \frac{1}{2} \left(\nabla_\mu \beta_\nu + \nabla_\nu \beta_\mu \right) \left[T^{\mu\nu}(\boldsymbol{\rho}_{12}) - T^{\mu\nu}(\tilde{\boldsymbol{\rho}}_{12}) \right] \\
&- (\partial_\mu \alpha) \left[N^\mu(\boldsymbol{\rho}_{12}) - N^\mu(\tilde{\boldsymbol{\rho}}_{12}) \right] \\
={}& -\mathfrak{s}(\boldsymbol{\rho}_{12}) + \nabla_\mu w^\mu(\tilde{\boldsymbol{\rho}}_{12}) \\
&- (\nabla_\mu \beta_\nu) T^{\mu\nu}(\boldsymbol{\rho}_{12}) - (\partial_\mu \alpha) N^\mu(\boldsymbol{\rho}_{12}),
\end{aligned} \tag{7.47}$$

giving us a generalization of (7.35).

BATH FLUID WITH GENERAL EXPONENTIAL FORM   At last, we study a rather general class of density operators, which we call general exponential density operators. They generalize global and local equilibrium states, but are still of exponential form

$$\tilde{\rho}_{12} = \frac{1}{Z} \exp\left[\int_{\Sigma(\tau)} d\Sigma_\mu \left(-h^\mu_{\alpha\beta} T^{\alpha\beta} - l^\mu_\alpha N^\alpha \right) \right] \tag{7.48}$$

with a nonequilibrium partition function

$$Z = \mathrm{Tr}\left\{ \exp\left[\int_{\Sigma(\tau)} d\Sigma_\mu \left(-h^\mu_{\alpha\beta} T^{\alpha\beta} - l^\mu_\alpha N^\alpha \right) \right] \right\}. \tag{7.49}$$

Here, $h^\mu_{\alpha\beta}(x)$ and $l^\beta_\alpha(x)$ are local parameter fields, which can be understood as Lagrange multipliers which enforce unrestricted expectation values $T^{\mu\nu}(\tilde{\rho}_{12})$ and $N^\mu(\tilde{\rho}_{12})$. In particular, the expectation values do not need to coincide with those in local or global equilibrium.

It is important to note that for a given hypersurface $\Sigma(\tau)$ with normal vector $n_\mu(x)$, only the contractions $n_\mu(x)h^\mu_{\alpha\beta}(x)$ and $n_\mu(x)l^\mu_\alpha(x)$ enter in the last equation. Hence, the components orthogonal to $n^\mu(x)$ can be changed without changing the state $\tilde{\rho}_{12}$ given in (7.48). In this sense, the degrees of freedom in the energy-momentum tensor and particle number current match with those of the parameter fields.

As before, the exponential form of (7.48) allows for a simple decomposition of the von Neumann entropy

$$S(\tilde{\rho}_{12}) = \ln Z + \int_{\Sigma(\tau)} d\Sigma_\mu \left(h^\mu_{\alpha\beta} T^{\alpha\beta}(\tilde{\rho}_{12}) + l^\mu_\alpha N^\alpha(\tilde{\rho}_{12}) \right). \tag{7.50}$$

If the Schwinger function $W = \ln Z$ is local (which has to be tested), one may define a nonequilibrium local entropy current via

$$s^\mu(\tilde{\rho}_{12}) = -h^\mu_{\alpha\beta} T^{\alpha\beta}(\tilde{\rho}_{12}) - l^\mu_\alpha N^\alpha(\tilde{\rho}_{12}) + w^\mu(\tilde{\rho}_{12}), \tag{7.51}$$

with (7.25) implied.

For the relative entropy between an arbitrary state ρ_{12} and a reference state $\tilde{\rho}_{12}$ of the form (7.48) we then find

$$\begin{aligned}
&S(\rho_{12}\|\tilde{\rho}_{12}) \\
&= -S(\rho_{12}) + \ln Z + \int_{\Sigma(\tau)} d\Sigma_\mu \left(h^\mu_{\alpha\beta} T^{\alpha\beta}(\rho_{12}) + l^\mu_\alpha N^\alpha(\rho_{12}) \right) \\
&= -S(\rho_{12}) + \int_{\Sigma(\tau)} d\Sigma_\mu \Big(-s^\mu(\tilde{\rho}_{12}) \\
&\quad + h^\mu_{\alpha\beta} \left[T^{\alpha\beta}(\rho_{12}) - T^{\alpha\beta}(\tilde{\rho}_{12}) \right] + l^\mu_\alpha \left[N^\alpha(\rho_{12}) - N^\alpha(\tilde{\rho}_{12}) \right] \Big).
\end{aligned} \tag{7.52}$$

The time evolution from $\Sigma(\tau_i)$ to $\Sigma(\tau_f)$ generated by the map $\mathcal{N}$ is constrained by the generalized second law in form of (7.30). Similar to before, one may construct a corresponding local form with respect to the model density operator $\tilde{\rho}_{12}$ using the latter identity. We omit the corresponding inequalities for brevity.

Part III

UNCERTAINTY

We propose a formulation of the uncertainty principle in terms of relative entropy, allowing us to describe discrete and continuous variable quantum systems (Chapter 8) as well as quantum fields consistently (Chapter 9). Further, we investigate entropic uncertainty in phase space (Chapter 10).

8 | REUR FOR QUANTUM SYSTEMS

The following discussion was published in [E]. S. F. proposed and supervised the project. B. H. was involved in the early developments of an EUR formulated with relative entropy in the form of a bachelor thesis [278]. However, none of the results presented in [E] are contained in [278]. The REUR was found by me and I wrote early versions of the draft. S. F. and I finalized the manuscript.

As anticipated in Section 3.5, there exist various EURs for discrete and continuous variables. However, the two cases of variables have only been unified in a rather formal way in [33] (see also the related works [123, 137, 144]), with the discrete MU relation (3.86) and the continuous FL relation (3.82) as special cases. The difficulties in such a unification can be traced back to the inconsistency of the Shannon entropy $S(p)$ and the differential entropy $S(f)$ in the continuum limit (2.12). We reformulate the results of [33] in terms of relative entropies, to obtain a relation which is capable of describing both types of variables at once. Besides being an interesting task by itself, this paves the ground for quantum fields, see Chapter 9.

Although the formulation of the uncertainty principle in terms of relative entropy has not been investigated so far, relative entropy is important in the context of entropic uncertainty. First, the monotonicity property (3.18) of the quantum relative entropy allows for a simple proof of the discrete MU relation (3.86) (see [12, 133, 279]). Second, the uncertainty principle can be formulated in terms of mutual information (which is defined as a relative entropy) in the presence of classical or quantum memory [136, 146, 147]. These relations state that the measurement outcomes of two non-commuting observables can not be arbitrarily correlated with a memory system and are consequently called information exclusion principles. Third, measurement EURs, which describe the disturbance of an observable caused by a previous measurement of another observable, have been formulated in terms of relative entropy in [280, 281].

The main idea is to consider reference distributions of maximum entropy, allowing for a decomposition of the relative entropy (similar to e.g. (6.24)). We discuss various choices explicitly for discrete variables in Section 8.1 and for continuous variables in Section 8.2. These considerations are unified in a single relation, the REUR, in Section 8.3. As the logic of this relation is in some sense reversed, we discuss it at length. Finally, we discuss two examples in Section 8.4.

8.1 DISCRETE VARIABLES

NO PRIOR INFORMATION Let us consider a discrete quantum system with $D < \infty$ and two discrete observables with operators $\mathbf{Y}$ and $\mathbf{Z}$. Measuring them leads to the two distributions $p(y)$ and $q(z)$ as defined in (3.70). A formulation of the uncertainty principle in terms of relative entropy mainly requires a specification of two reference distributions $\tilde{p}(y)$ and $\tilde{q}(z)$. We will see that suitable reference distributions are given by Boltzmann-type distributions or more generally by distributions maximizing an entropy. In this way, additional information about the distributions of interest will be encoded directly in the resulting uncertainty relation.

We start with the simplest case where no prior information about $p(y)$ and $q(z)$ is available. A suitable reference distribution is picked with the maximum entropy principle (or equivalently the principle of minimum expected relative entropy), which postulates the uniform distribution

$$\tilde{p}(y) = p_{\max}(y) = \frac{1}{D}. \tag{8.1}$$

The corresponding relative entropy $S(p\|p_{\max})$ simplifies to

$$S(p\|p_{\max}) = -S(p) + S(p_{\max}) = -S(p) + \ln D, \tag{8.2}$$

such that the MU relation (3.86) becomes

$$S(p\|p_{\max}) + S(q\|q_{\max}) \leq \ln c - S(\boldsymbol{\rho}) + 2\ln D. \tag{8.3}$$

Interestingly, a sum of relative entropies is bounded from above in a non-trivial way: the trivial bound $2\ln D$, which is a consequence of

$$S(p\|p_{\max}) + S(q\|q_{\max}) \leq S(p_{\max}) + S(q_{\max}) = 2\ln D, \tag{8.4}$$

is reduced by the quantum incompatibility measure c (note here that $\ln c \leq 0$) and the mixedness of the quantum state $S(\boldsymbol{\rho})$.

To exemplify the information theoretic significance, let us consider the special case where $\mathbf{Y}$ and $\mathbf{Z}$ are mutually unbiased such that perfect knowledge about one observable implies no knowledge about the other. In this case we have

$$c = |\langle y|z\rangle|^2 = \frac{1}{D} \tag{8.5}$$

for all y and z, such that $\ln c = -\ln D$. Further, we consider the state to be maximally mixed with $S(\boldsymbol{\rho}) = \ln D$. Then, the right hand side of (8.3) reduces to zero and hence $p(y)$ and $q(z)$ become indistinguishable from uniform distributions with maximum entropy.

GIVEN EXPECTATION VALUES   Next, we consider a situation where we have access to a set of expectation values. Note that this includes the possibility of measuring other observables, as long as their measurement outcomes are distributed according to the same probability distributions. This is fulfilled in particular if the two observables commute as a consequence of the definition (3.70).

For example, we may consider the mean value of Y,

$$\mu_y = \sum_y p(y)y, \tag{8.6}$$

such that $\tilde{p}(y)$ is required to have the same mean value, i.e.

$$\tilde{\mu}_y = \sum_y \tilde{p}(y)y \equiv \mu_y. \tag{8.7}$$

The associated maximum entropy distribution is chosen as reference distribution

$$\tilde{p}(y) = p_{\max}(y) = \frac{1}{Z_y}e^{-\gamma_y y}, \tag{8.8}$$

with γ_y being a Lagrange multiplier and Z_y denoting a normalization constant.

The mean values of other commuting observables can be included in the same way. An interesting choice can be made when the Hamiltonian H commutes with Y, i.e. $[H, Y] = 0$. Then, one may measure the energy expectation value to obtain a thermal reference distribution $\tilde{p}(y) = p_{\max}(y)$, which is uniquely determined by the optimal inverse temperature β_y.

For the Boltzmann-type reference distribution (8.8) the relative entropy becomes

$$S(p\|p_{\max}) = -S(p) + S(p_{\max}) = -S(p) + \ln Z_y + \gamma_y\mu_y. \tag{8.9}$$

Hence, a reformulation of the MU EUR (3.86) reads

$$S(p\|p_{\max}) + S(q\|q_{\max}) \leq \ln c - S(\boldsymbol{\rho}) + \ln(Z_yZ_z) + \gamma_y\mu_y + \gamma_z\mu_z. \tag{8.10}$$

Note that the additional quantities appearing in the bound can be computed from the true distributions $p(y)$ and $q(z)$, such that they can be fully determined after measuring $p(y)$ and $q(z)$.

One can generalize these considerations and choose a maximum entropy reference distribution $\tilde{p}(x)$ such that the measured expectation value μ_y does *not* agree with $\tilde{\mu}_y$. In this case, the relative entropy becomes

$$S(p\|p_{\max}) = -S(p) + S(p_{\max}) + \gamma_y(\mu_y - \tilde{\mu}_y). \tag{8.11}$$

However, as the additional term is also contained in the maximum entropy expression $S(p_{\max})$, the relation (8.10) remains unmodified.

GIVEN SET OF MOMENTS   We go one step further and allow for an arbitrary set of moments to be constrained. More precisely, we fix the M expectation values

$$\langle m_j \rangle = \sum_y p(y) m_j(y), \tag{8.12}$$

with $m_1(y), ..., m_M(y)$ being polynomials in y. Then, the reference distribution needs to fulfill the set of constraints

$$\alpha_j(p, \tilde{p}) = \sum_y \left(\tilde{p}(y) - p(y) \right) m_j(y) = 0 \tag{8.13}$$

for all $j \in \{1, ..., M\}$. If a solution to this optimization problem exists, which is typically the case for discrete, but not for continuous variables, it can be computed from the maximum entropy principle

$$\tilde{p}(y) = p_{\max}(y) = \exp\left(\sum_{j=0}^{M} \lambda_j m_j(y) \right). \tag{8.14}$$

Therein, λ_0 ensures normalization, while λ_j with $j > 0$ implement the M conditions $\alpha_j(p, \tilde{p}) = 0$. Note that the form of this solution is the same for constraints given in terms of inequalities $\alpha_j(p, \tilde{p}) \geq 0$, requiring additional constraints $\lambda_j \geq 0$ for all $j \in \{1, ..., M\}$ [228, 229].

FORMULATION SOLELY IN TERMS OF RELATIVE ENTROPIES   At last, we consider another rewriting of the MU relation (3.86). In particular, we also want the quantum entropy $S(\rho)$ to be replaced by a suitable quantum relative entropy $S(\rho \| \tilde{\rho})$. To that end, we consider maximum entropy states by starting from the maximum entropy distribution $p_{\max}(y)$ and constructing the corresponding measured quantum state $\rho_{Y,\max}$ via

$$\rho_{Y,\max} = \sum_y p_{\max}(y) |y\rangle \langle y|. \tag{8.15}$$

We then find for the quantum relative entropy between the state ρ and such a reference state

$$\begin{aligned} S(\rho \| \tilde{\rho}) = S(\rho \| \rho_{Y,\max}) &= -S(\rho) + S(\rho_{Y,\max}) \\ &= -S(\rho) + S(p_{\max}). \end{aligned} \tag{8.16}$$

Moreover, if one chooses the maximally mixed state

$$\rho_{\max} = \frac{1}{D} \mathbb{1} \tag{8.17}$$

as a reference distribution, the corresponding quantum relative entropy can be written as

$$S(\rho \| \tilde{\rho}) = S(\rho \| \rho_{\max}) = -S(\rho) + \ln D. \tag{8.18}$$

Combining the latter equations leads to a reformulation of (3.86) solely in terms of relative entropies

$$
\begin{aligned}
S(p\|p_{\max}) &+ S(q\|q_{\max}) \\
&\leq \ln(cD) - S(\boldsymbol{\rho}\|\boldsymbol{\rho}_{\max}) + S(\boldsymbol{\rho}\|\boldsymbol{\rho}_{Y,\max}) + S(\boldsymbol{\rho}\|\boldsymbol{\rho}_{Z,\max}).
\end{aligned}
\tag{8.19}
$$

Note that the maximum overlap fulfills $c \geq \frac{1}{D}$ such that $\ln(cD) \geq 0$, with equality if and only if the measurement bases are mutually unbiased.

8.2 CONTINUOUS VARIABLES

NOTE ON MEASURABILITY For position X and momentum K, the overlap is given by (3.39), showing that

$$
c = |\langle x|k\rangle|^2 = \frac{1}{2\pi}
\tag{8.20}
$$

for all x, k, rendering the bases $\{|x\rangle\}_x$ and $\{|k\rangle\}_k$ mutually unbiased in a continuous sense.

However, it is important to note that one can not measure continuous outcomes in practice. More precisely, any actual experiment is only capable of measuring X and K with finite precision Δx and Δk, respectively. For example, one ends up with a discrete probability distribution [113]

$$
p(x) = \int_{(x-1/2)\,\Delta x}^{(x+1/2)\Delta x} \mathrm{d}x'\, f(x').
\tag{8.21}
$$

describing the probability to measure the position somewhere in the interval $[(x - 1/2)\Delta x, (x + 1/2)\Delta x]$.

Unfortunately, the limit of infinite precision $\Delta x \to 0$ leads to a diverging Shannon entropy $S(p)$ as described in Section 2.2. Such problems might be avoided when working in terms of relative entropies. We discuss the effects of discretization in more detail in Section 13.3.

GAUSSIAN DISTRIBUTIONS AS OPTIMAL MODELS Also for continuous variables one could consider a given mean value

$$
\mu_x = \int \mathrm{d}x\, f(x)\, x
\tag{8.22}
$$

as side constraint to obtain a suitable reference distribution, as investigated in Section 8.1. Furthermore, if the quantum system of interest has a finite length, e.g. $x \in [-L/2, L/2]$, one may choose the uniform reference distribution $\tilde{f}(x) = f_{\max}(x) = 1/L$ for position measurements if no constraints are given. Note that the uniform distribution may not be chosen for *both* continuous variables simultaneously as

$x \in [-L/2, L/2]$ still allows for (discrete) $k \in (-\infty, \infty)$, in which case the uniform distribution does not exist.

However, for continuous variables, a fixed variance

$$\sigma_x^2 = \int \mathrm{d}x \, f(x)(x - \mu_x)^2 \tag{8.23}$$

is of special interest, as the corresponding maximum entropy distribution is given by a Gaussian

$$\tilde{f}(x) = f_{\max}(x) = \frac{1}{\sqrt{2\pi\sigma_x^2}} e^{-\frac{1}{2}\left(\frac{x-\mu_x}{\sigma_x}\right)^2}. \tag{8.24}$$

Its differential entropy reads

$$S(f_{\max}) = \frac{1}{2} \ln \left(2\pi e \sigma_x^2 \right), \tag{8.25}$$

such that the FL EUR (3.82), which can be considered the continuous variable analog of the MU relation (3.86), becomes

$$S(f\|f_{\max}) + S(g\|g_{\max}) \leq -S(\rho) + 1 + \ln(\sigma_x \sigma_k). \tag{8.26}$$

Similar to the discrete case, the additional terms in the bound follow from the two distributions $f(x)$ and $g(k)$.

NOTE ON REWRITING WITH QUANTUM RELATIVE ENTROPY Unfortunately, we did not find a way to rewrite the latter relation solely in terms of relative entropy, which is due to the non-existence of a post-measurement state for measurements of continuous variables (cf. Section 3.4). Hence, a similar line of reasoning as presented in Section 8.1 fails.

NOTE ON REWRITING THE BBM RELATION Instead of rewriting the FL relation (3.82) in terms of relative entropies with respect to Gaussian models, one may equally well consider the BBM EUR (3.79), which leads to a different upper bound

$$S(f\|f_{\max}) + S(g\|g_{\max}) \leq \ln 2 + \ln(\sigma_x \sigma_k). \tag{8.27}$$

This relation will be the starting point of formulating the concept of entropic uncertainty for quantum fields in Chapter 9. Also, this rewriting is particularly useful to understand in which sense the BBM relation is stronger than Heisenberg's relation (3.80). Solving for the variance deviation product gives

$$\sigma_x^2 \sigma_k^2 \geq \frac{1}{4} e^{2S(f\|f_{\max}) + 2S(g\|g_{\max})}, \tag{8.28}$$

showing that the BBM relation is stronger than the Heisenberg relation if and only if the distributions $f(x)$ and $g(k)$ are of non-Gaussian form, with non-Gaussianity measured in terms of relative entropies (see also [29, 282, 283]). However, the relation (8.27) holds only for continuous variables and thus is not suitable for formulating a relation being capable of describing discrete as well as continuous variables simultaneously.

8.3 RELATIVE ENTROPIC UNCERTAINTY RELATION – PART I

MODEL DISTRIBUTIONS OF MAXIMUM ENTROPY Having discussed the basic strategy from various perspectives, let us now formalize our considerations to obtain our desired unified relation. Consequently, we use the expressions $p(y)$ and $q(z)$ for *both* types of distributions, i.e. discrete as well as continuous, in the following.

We have seen that it is useful to pick the reference distributions $\tilde{p}(y)$ and $\tilde{q}(z)$ such that the relative entropies $S(p\|\tilde{p})$ and $S(q\|\tilde{q})$ become differences of entropies. It is simple to show that this occurs if and only if they correspond to maximum entropy distributions.

Let us start from the definition of the classical relative entropy

$$S(p\|\tilde{p}) = -S(p) + S(p, \tilde{p}), \tag{8.29}$$

where the second term is the so-called cross entropy

$$S(p, \tilde{p}) = -\sum_y p(y) \ln \tilde{p}(y). \tag{8.30}$$

Note that for continuous distributions the sum is replaced by an integral. The reference distribution $\tilde{p}(y)$ should be chosen such that

$$S(p, \tilde{p}) = S(\tilde{p}) \tag{8.31}$$

for a set of constraints

$$\alpha_j(p, \tilde{p}) = 0 \tag{8.32}$$

with $j \in \{1, ..., M\}$ where $M \in \mathbb{N}$. Examples for these constraints have been discussed in detail in the two previous sections.

Plugging the condition (8.31) into the definition of the relative entropy (8.29) together with non-negativity of the latter yields

$$S(p\|\tilde{p}) = -S(p) + S(\tilde{p}) \geq 0 \Leftrightarrow S(p) \leq S(\tilde{p}) \tag{8.33}$$

for all allowed distributions $p(y)$. Hence, $\tilde{p}(y)$ does necessarily correspond to the maximum entropy distribution $\tilde{p}(y) = p_{\max}(y)$ under the constraints (8.32). In this sense, this choice is *optimal*.

Note that if the condition (8.31) is released such that terms linear in the constraints are allowed, we still obtain a maximum entropy distribution. However, it is less optimal as not all of its properties agree with those of the true distribution. For example, their mean values could differ, i.e. $\mu_y \neq \tilde{\mu}_y$.

Furthermore, optimal model distributions do often fulfill the support condition $\operatorname{supp}[p(y)] \subseteq \operatorname{supp}[p_{\max}(y)]$ for all y, such that the corresponding relative entropy $S(p\|\tilde{p})$ remains finite. To be more precise, the support condition is always fulfilled for discrete variables. For continuous variables, it can be violated. We will further elaborate on this issue when discussing the quantum origin of the resulting bound on a sum of two relative entropies.

The relative entropy decomposes linearly whenever the reference distribution maximizes an entropy.

UNCERTAINTY RELATION AND DISCUSSION   Without further ado, we state the REUR for two (discrete or continuous) distributions $p(y)$ and $q(z)$ with respect to maximum entropy reference distribution $p_{max}(y)$ and $q_{max}(z)$ under some constraints

$$S(p\|p_{max}) + S(q\|q_{max}) \leq \ln c - S(\boldsymbol{\rho}) + S(p_{max}) + S(q_{max}), \quad (8.34)$$

where c still denotes the maximum overlap between any two eigenstates for PVMs (3.87) or its generalization for POVMs (3.88). As it corresponds to a reformulation of the measure-theoretic formulation in [33], we omit a proof for brevity. The main ingredients in a proof are the Golden-Thompson inequality and Gibbs variational principle.

To better understand how this relation expresses the uncertainty principle, let us discuss it in detail.

a) *Sum of relative entropies is bounded from above.*

A sum of two relative entropies is trivially bounded from below by zero as relative entropy is non-negative. The non-trivial bound expressing uncertainty comes from above and not from below. This shows that the joint distinguishability of $p(y)$ and $q(z)$ with respect to reference distributions of minimum information content is bounded.

In general, it requires more information about a distribution $p(y)$ additional to the information encoded in the maximum entropy reference $\tilde{p}(y)$ to distinguish the two. This surplus of information is measured in terms of relative entropy and bounded from above, which is in accordance with the usual logic of the uncertainty principle, i.e. the fact that one can never have the full information about two non-commuting observables.

b) *Bound is of quantum origin.*

At first, note that $\ln c - S(\boldsymbol{\rho}) \leq 0$. Hence, a classical upper bound can be obtained by considering the absence of missing information about the state $S(\boldsymbol{\rho}) = 0$ and two measurement bases which share at least one eigenvector such that $\ln c = 0$. This leads to

$$S(p\|p_{max}) + S(q\|q_{max}) \leq S(p_{max}) + S(q_{max}). \quad (8.35)$$

For discrete variables, the largest classical upper bound

$$S(p_{max}) + S(q_{max}) \leq 2 \ln D \quad (8.36)$$

is finite for a finite Hilbert space dimension $D < \infty$. This bound gets reduced (by a finite amount) by the uncertainty principle in terms of the quantum incompatibility c and the state's mixedness $S(\boldsymbol{\rho})$ impeding to distinguish the distributions from each other.

For continuous variables, the Hilbert space dimension is infinite $D = \infty$ and hence the classical upper bound $S(p_{max}) + S(q_{max})$

may diverge to $+\infty$. For example, considering one Gaussian reference distribution for a fixed variance in the limit of arbitrarily large width leads to a diverging entropy. Nevertheless, this (possibly infinite) classical bound gets still reduced by $\ln c - S(\boldsymbol{\rho})$.

Hence, we can conclude that in both cases the uncertainty principle provides a smaller bound than any classical bound showing its quantum origin. This is in accordance with the corresponding EURs (3.86) and (3.82), whose minimum classical lower bounds are given by 0 and $-\infty$, respectively. However, they encode the uncertainty principle by lifting a lower bound instead of reducing an upper bound.

c) *MU and FL relations are special cases.*

As already discussed in the previous two sections, our REUR given in (8.34) reproduces the MU relation (3.86) for discrete and the FL relation (3.82) for continuous variables, respectively. Consequently, its tightness depends on the type of variables. For discrete variables, equality is achieved for states being diagonal in one of the two bases which have to be mutually unbiased. For continuous variables, the relation is tight in the infinite temperature limit.

d) *Discrete and continuous spectra are unified.*

In contrast to all other existing EURs, the REUR is simultaneously well-defined for discrete as well as continuous variables, which is one of the main advantages relative entropy has over entropy. Note that the entropies appearing in the bound can always be expressed in terms of the constrained quantities after plugging in the corresponding expressions for $p_{\max}(y)$ and $q_{\max}(z)$, rendering the bound free of ambiguities. Also, the bound in (8.34) is indeed well-behaved in the continuum limit as all possible infinities appearing in c and $S(p_{\max}) + S(q_{\max})$ cancel out.

e) *Bound contains side information.*

Again in contrast to other EURs, the bound in the REUR (8.34) contains side information in terms of the reference distributions $p_{\max}(y)$ and $q_{\max}(z)$, which is interesting for experimental applications.

f) *Invariance under change of normalization.*

Another advantage relative entropy has over entropy is its invariance under scaling transformations, which immediately carries over to the left hand side of the REUR (8.34). For the right hand side, we consider a transformation $|y\rangle \rightarrow |y'\rangle = a|y\rangle$ with $a \in \mathbb{R}$ and $|a| \neq 1$, which corresponds to a change of normalization of the basis $\{|y\rangle\}_y$. The integration measure transforms as $\mathrm{d}y \rightarrow \mathrm{d}y' = a^{-2}\mathrm{d}y$ such that the probability $p(y)\mathrm{d}y \rightarrow$

$p(y')dy' = p(y)dy$ remains unchanged. However, the probability density scales as $p(y) \rightarrow p(y') = a^2 p(y)$ and hence the quantities in the bound transform according to

$$\ln c \rightarrow \ln c + \ln a^2, \quad S(p_{\max}) \rightarrow S(p_{\max}) - \ln a^2, \quad (8.37)$$

showing that also the bound remains invariant under scaling transformations.

8.4 EXAMPLES

SPIN–1 OBSERVABLES To illustrate the logic of the REUR (8.34) we consider spin observables J_x, J_y, J_z, which fulfill a $SU(2)$ algebra

$$[J_k, J_l] = i\epsilon_{klm} J_m, \quad (8.38)$$

where $k, l, m \in \{x, y, z\}$ label spatial dimensions and ϵ_{klm} denotes the totally antisymmetric Levi-Civita tensor. The spin operators fulfill eigenvalues equations of the form

$$J_k \, |j_k\rangle = j_k \, |j_k\rangle \,, \quad (8.39)$$

where $j_k \in \{-1, 0, 1\}$ denote the eigenvalues and $|j_k\rangle$ the corresponding normalized eigenvectors.

We restrict our analysis to a fixed total spin of $J = 1$, in which case the spin operators live in the adjoint representation of $SU(2)$. Then, the three sets of eigenvectors span orthonormal bases for the 3-dimensional Hilbert space $\mathcal{H}$ from which one can construct PVMs, each describing the measurement of one spin component J_k. When measuring two spatial components simultaneously, we are limited by the uncertainty principle as a result of the non-trivial commutation relation (8.38).

Without loss of generality we work in the basis spanned by $\{|j_z\rangle\}_{j_z}$, which we denote with $\{|-1\rangle, |0\rangle, |1\rangle\}$. In $\mathbb{R}^3$ the state vectors are conveniently represented as unit vectors

$$|1\rangle \cong \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad |0\rangle \cong \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad |-1\rangle \cong \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \quad (8.40)$$

In this basis, J_z becomes diagonal and the spin operators are represented by hermitian 3×3 matrices

$$J_x = \frac{1}{\sqrt{2}} \left(|1\rangle \langle 0| + |0\rangle \langle 1| + |0\rangle \langle -1| + |-1\rangle \langle 0| \right)$$

$$\hat{=} \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix},$$

$$J_y = \frac{i}{\sqrt{2}} \left(-|1\rangle \langle 0| + |0\rangle \langle 1| - |0\rangle \langle -1| + |-1\rangle \langle 0| \right)$$

$$\hat{=} \frac{i}{\sqrt{2}} \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \qquad (8.41)$$

$$J_z = |1\rangle \langle 1| + |-1\rangle \langle -1| \hat{=} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

The eigenvectors of the J_x and J_y operators are given by (ordered corresponding to the eigenvalues $1, 0, -1$, respectively)

$$\{j_x\}_{j_x} = \left\{ \frac{1}{2} \left(|1\rangle + \sqrt{2}|0\rangle + |-1\rangle \right), \frac{1}{\sqrt{2}} \left(|1\rangle - |-1\rangle \right), \right.$$

$$\left. \frac{1}{2} \left(|1\rangle - \sqrt{2}|0\rangle + |-1\rangle \right) \right\},$$

$$\{j_y\}_{j_y} = \left\{ \frac{1}{2} \left(-|1\rangle - i\sqrt{2}|0\rangle + |-1\rangle \right), \frac{1}{\sqrt{2}} \left(|1\rangle + |-1\rangle \right), \right.$$

$$\left. \frac{1}{2} \left(-|1\rangle + i\sqrt{2}|0\rangle + |-1\rangle \right) \right\}, \qquad (8.42)$$

showing that the corresponding measurement bases are not mutually unbiased. Note that only the Pauli matrices σ_k, which live in the fundamental representation of $SU(2)$, have mutually unbiased eigenvectors.

We want to measure J_y and J_z, in which case we are limited by the quantum incompatibility (3.87), which evaluates to

$$c = \max \left\{ 0, \frac{1}{4}, \frac{1}{2} \right\} = \frac{1}{2}. \qquad (8.43)$$

The entropies of the measured distributions $f(j_y) = \langle j_y | \rho | j_y \rangle$ and $g(j_z) = \langle j_z | \rho | j_z \rangle$ are bounded from below by the MU relation (3.86)

$$S(p) + S(q) \geq \ln 2 + S(\rho), \qquad (8.44)$$

while the relative entropies with respect to uniform model distributions are bounded from above by the REUR (8.34)

$$S(p\|p_{max}) + S(q\|q_{max}) \leq \ln \frac{9}{2} - S(\boldsymbol{\rho}). \tag{8.45}$$

Note that for discrete variables we also have

$$S(\boldsymbol{\rho}) \leq S(p), S(q). \tag{8.46}$$

Bounding an entropy from below is equivalent to bounding a relative entropy from above.

We consider now the case where the state is pure $S(\boldsymbol{\rho}) = 0$. Without loss of generality we assume that the measurement of $\boldsymbol{J}_y$ is never more uncertain than the measurement of $\boldsymbol{J}_z$, such that $S(p) \leq S(q)$. Then, for $S(p) \geq 0$ the uncertainty principle implies $S(q) \geq \ln \sqrt{2}$. Equivalently, we have $S(p\|p_{max}) \leq \ln 3$ and hence $S(q\|q_{max}) \leq \ln \frac{3}{\sqrt{2}}$. The equivalence of these two statements is illustrated in Figure 8.1.

(a) Allowed $p(j_y)$ (b) Allowed $q(j_z)$
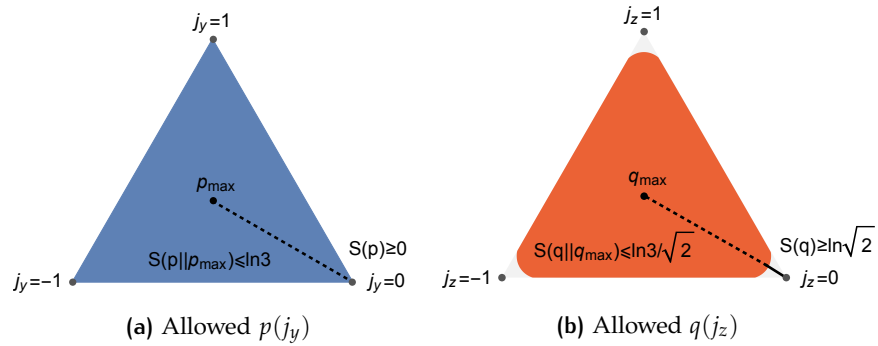
Figure 8.1: Allowed regions for the distributions $p(j_y)$ (blue area) in (a) and $q(j_z)$ (red area) in (b) describing $\boldsymbol{J}_y$ and $\boldsymbol{J}_z$ measurements in a spin-1 system as required by the uncertainty principle and $S(p) \leq S(q)$. While entropies (black solid lines) are bounded from below, relative entropies (black dotted lines) are bounded from above.

ANGLE AND ANGULAR MOMENTUM At last, we demonstrate the strengths of quantifying entropic uncertainty in terms of relative entropy by considering a particle on a ring of radius R, which is characterized by the z-component of its angular momentum $\boldsymbol{J}_z$ and its position in terms of an angle ϕ. In particular, we exemplify how the bound behaves when taking the continuum and infinite volume limits.

We consider again the angular momentum eigenstates $|j_z\rangle$ defined through (8.39). The corresponding Hilbert space has the dimension $D = 2J + 1$, where J denotes the total angular momentum, which is a half-integer in general. As the set $\{|j_z\rangle\}_{j_z}$ forms a orthonormal basis for $\mathcal{H}$, the set $\{|j_z\rangle \langle j_z|\}_{j_z}$ is a PVM.

The continuous angle states can be defined via [284]

$$|\phi\rangle = \frac{1}{\sqrt{2J+1}} \sum_{j_z=-J}^{J} e^{-ij_z\phi} |j_z\rangle, \tag{8.47}$$

which is consistent with periodicity $|\phi\rangle = |\phi + 2\pi\rangle$ or more generally that the angular momentum operator J_z generates rotations on the ring

$$e^{-i\varphi J_z} |\phi\rangle = |\phi + \varphi\rangle. \tag{8.48}$$

The angle states have a non-trivial overlap

$$\langle \varphi | \phi \rangle = \frac{1}{2J+1} \frac{\sin\left[\left(J + \frac{1}{2}\right)(\varphi - \phi)\right]}{\sin\left[\frac{1}{2}(\varphi - \phi)\right]} \tag{8.49}$$

and resolve the identity

$$\mathbb{1} = (2J+1) \int_0^{2\pi} \frac{d\phi}{2\pi} |\boldsymbol{\phi}\rangle \langle \boldsymbol{\phi}| = \sum_{j_z=-J}^{J} |j_z\rangle \langle j_z|. \tag{8.50}$$

Hence, they form an overcomplete basis similar to canonical coherent states introduced in Section 3.4, and the set $\{|\boldsymbol{\phi}\rangle \langle \boldsymbol{\phi}|\}_\phi$ forms a POVM. One can also consider a discrete subset of angle states $|\theta_j\rangle$

$$\theta_j = \theta_0 + \frac{2\pi j}{2J+1} \tag{8.51}$$

with $j \in \{0, ..., 2J\}$ and $\theta_0 \in [0, 2\pi)$ being a free parameter, which is typically set to zero. These angle states are chosen such that they are orthonormal $\langle \theta_j | \theta_k \rangle = \delta_{jk}$, which follows from (8.49), and hence their measurement corresponds to a PVM.

The limit $J \to \infty$ renders the discrete states dense and therefore can be considered as a continuum limit. However, the resulting angle states differ in normalization from the continuous angle states (8.47). Also, one may consider the infinite volume limit $R \to \infty$, which can be implemented after introducing positions $x = 2\pi R\phi$ and conjugate momenta $k = j_z/(2\pi R)$. In the following, we discuss the behavior of the REUR (8.34) for these two limits.

For finite J, the measurement of the angular momentum operator J_z leads to a discrete probability distribution $q(j_z)$. Hence, suitable reference distributions can be found by fixing moments of this distribution as discussed in Section 8.1. For example, one may take the uniform distribution $q_{max}(j_z) = \frac{1}{D}$ if no additional information is available, or, for known $\langle J_z \rangle$ and $\langle J_z^2 \rangle$ one may use

$$q_{max}(j_z) = \exp\left(\lambda_0 + \lambda_1 j_z + \lambda_2 j_z^2\right), \tag{8.52}$$

with Lagrange multipliers λ_j. To compute the contribution to the bound of the REUR, we evaluate the entropy of such a distribution

$$S(q_{max}) = -\lambda_0 - \lambda_1 \langle J_z \rangle - \lambda_2 \langle J_z^2 \rangle. \tag{8.53}$$

In the continuum limit $J \to \infty$, the distribution (8.52) is normalizable as long as $\lambda_2 < 0$. In this case, it corresponds to a Gaussian-like

distribution, which becomes continuous if one also takes the infinite volume limit $R \to \infty$.

For the discrete angle states (8.51) and finite J the situation is completely analogous to the angular momentum eigenstates. Also, the overlap evaluates to

$$c = |\langle j_z | \theta_j \rangle|^2 = \frac{1}{2J+1} = \frac{1}{D} \tag{8.54}$$

for all j_z, θ_j showing that the two measurement bases are mutually unbiased. Exemplary, we depict such a measurement in Figure 8.2a for $J = 1/2$.

The measurement of the continuous angles ϕ requires a more careful analysis. If no additional information is available, the (circular) uniform distribution $p_{\max}(\phi) = \frac{1}{2\pi}$ is still a reasonable reference distribution as the angle ϕ is drawn from a finite range $\phi \in [0, 2\pi)$. The corresponding entropy is $S(p_{\max}) = \ln(2\pi)$, which corresponds to a global maximum.

It is also interesting to consider side information about moments. In directional statistics, which deals with probability distributions on spheres, it is convenient to introduce the so-called circular moments, which are defined as the expectation values of powers of the complex-valued variable $e^{i\phi}$ [285]. Fixing the first circular moment $\langle e^{i\phi} \rangle$ leads to the von Mises distribution

$$p_{\max}(\phi) = \frac{e^{\kappa \cos(\phi - \mu)}}{2\pi I_0(\kappa)}, \tag{8.55}$$

where $\mu \in [0, 2\pi)$ and $\kappa \in \mathbb{R}^+$ are similar to mean and the inverse variance for a Gaussian distribution, while $I_j(\kappa)$ denotes the modified Bessel function of first kind of order j. Note that the von Mises distribution reduces to the uniform distribution in the limit $\kappa \to 0$.

The first circular moment evaluates to

$$\langle e^{i\phi} \rangle = \frac{I_1(\kappa)}{I_0(\kappa)} e^{i\mu}, \tag{8.56}$$

which can be used to fix μ and κ as these quantities are encoded in the phase and the magnitude of the latter expression, respectively.

Then, we can compute the entropy, which gives

$$S(p_{\max}) = \ln \left[2\pi I_0(\kappa) \right] - \kappa \frac{I_1(\kappa)}{I_0(\kappa)}. \tag{8.57}$$

Finally, the quantum incompatibility, now computed through (3.88), is given by $c = 1/(2\pi)$, such that the bound is fully determined. A measurement of continuous angles is shown in Figure 8.2b.

If one would have started with the discrete angles θ_j and taken the continuum limit $J \to \infty$ instead, $S(p_{\max})$ would have acquired a (diverging) negative offset $\ln \frac{2\pi}{2J+1}$, which would have been canceled

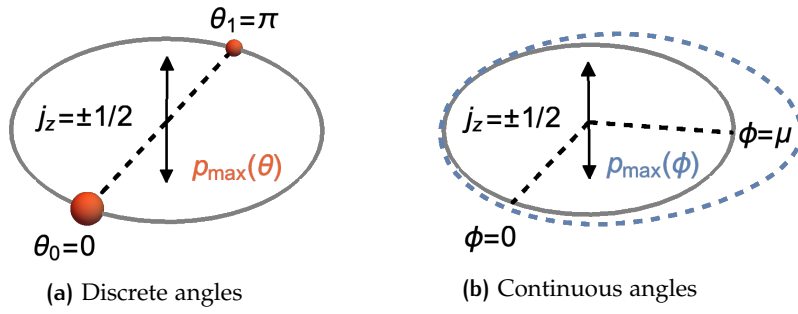**(a)** Discrete angles

(b) Continuous angles

Figure 8.2: Discrete angles $\theta_j = \theta_0 + \pi j$ with $j \in \{0,1\}$ in (a) versus continuous angles ϕ in (b) for a total spin $J = \frac{1}{2}$. While for the discrete angular momenta j_z we do not implement any side information, i.e. $q_{max}(j_z) = \frac{1}{2}$, we have chosen reference distributions which overweight certain angles. In (a) we have $p_{max}(\theta_0) > p_{max}(\theta_1)$ by fixing $\langle J_z \rangle$ (indicated by the size of the red dots) and in (b) we have chosen a von Mises distribution (8.55) with fixed $\mu = \frac{\pi}{2}$ and $\kappa = 2$ (indicated by the distance of the blue dashed line to the ring).

out by an analogous term in the quantum incompatibility $\ln c$. This shows again that the bound is well-behaved in the continuum limit.

Finally, let us mention that if κ becomes large, the von Mises distribution approaches a Gaussian distribution around $\phi = \mu$ with variance $1/\kappa$. If one additionally takes the infinite volume limit $R \to \infty$, both reference distributions are optimal for a given variance (see also discussion in Section 8.2).

9

REUR FOR QUANTUM FIELDS

This chapter is taken from [B]. S. F. proposed and supervised the project. The field-theoretic REUR was discovered by me. M. S. and I did the calculations. I wrote large parts of the manuscript and produced the figures. All authors contributed in finalizing the draft.

Information theoretic concepts have become more and more important for a deeper understanding of QFTs. Examples include intriguing phenomena such as entanglement [193–195], thermalization [G, 187, 191] and black holes [175–177]. However, a rigorous information theoretic treatment of the uncertainty principle has been missing so far.

Relative entropy turned out to be useful for formulating EURs in a rather universal way. In particular, it allowed us to describe discrete and continuous variables in a single relation. In this chapter, we go one step further and extend the concept of entropic uncertainty to quantum fields. Due to divergences in the field theory limit (not to be confused with the UV-divergences of the entanglement entropy (4.22)), no EUR for a QFT has been formulated so far. We overcome this problem by formulating the BBM relation in terms of relative entropies.

Another motivation is the close connection between uncertainty relations and entanglement witnesses (see discussion in Section 3.6 and also Chapter 12 and Chapter 13). However, describing entanglement in field theories accurately is a rather hard task. In particular, we have seen that the entanglement entropy is UV-divergent in the field theory limit (cf. Section 4.2), showing that one needs to follow a different line of reasoning compared to finite quantum systems. In this sense, the field-theoretic REUR paves the ground for well-defined entropic entanglement witnesses in QFTs.

We will develop the field-theoretic REUR for a free scalar quantum field. In this simple case the commutation relation of the field operators is bosonic (in this sense it is the field-theoretic extension of the Heisenberg algebra between position and momentum operators) and the vacuum state is of Gaussian form. Extensions to interacting theories are expected to be possible in the perturbative regime, although the bound of the relation might be modified by additional terms.

We start with an introduction to the Schrödinger functional formalism in Section 9.1. Then, we show that functional entropies are ill-defined and thereupon derive our field-theoretic REUR in Section 9.2. We discuss several examples and the connection to the multi-dimensional Heisenberg relation in Section 9.3, where we also comment on the measurability of the involved distributions.

9.1 FUNCTIONAL PROBABILITY DENSITIES

FROM OSCILLATOR MODES TO A QUANTUM FIELD   To make the dependence of all quantities on the number of modes explicit, we start from a set of coupled harmonic oscillators in $d = 1 + 1$ spacetime dimensions (note that the whole argument can be generalized easily to more than one spatial dimensions) with Hamiltonian

$$H = \frac{1}{2} \sum_j \epsilon \left[\pi_j^2 + \frac{1}{\epsilon^2} \left(\phi_j - \phi_{j-1} \right)^2 + m^2 \phi_j^2 \right]. \tag{9.1}$$

Therein, the $N \in \mathbb{N}$ modes are labeled by the index $j \in \{1, ..., N\}$ and carry the two real degrees of freedom ϕ_j and π_j. We assume periodic boundary conditions $\phi_0 = \phi_N$, connecting the chain of oscillators to a ring. The distance between two neighboring modes is given by the lattice constant ϵ. The continuum limit is given by $\epsilon \to 0$ (with the effect that the UV-regulator $1/\epsilon$ is released), while the oscillator picture can be emphasized with the choice $\epsilon = 1$.

To diagonalize the Hamiltonian (9.1), we use a discrete Fourier transformation

$$\phi_j = \sum_l \frac{\Delta k}{2\pi} e^{i\Delta kl\epsilon j} \tilde{\phi}_l, \quad \pi_j = \sum_l \frac{\Delta k}{2\pi} e^{-i\Delta kl\epsilon j} \tilde{\pi}_l, \tag{9.2}$$

where we introduced a new integer-valued index $-\frac{N}{2} \leq l < \frac{N}{2}$ for the momentum modes. As usual, we have the relations $\tilde{\phi}_l^* = \tilde{\phi}_{-l}$ and $\tilde{\pi}_l^* = \tilde{\pi}_{-l}$ as both fields are complex-valued in momentum space. The corresponding lattice constant reads $\frac{\Delta k}{2\pi}$, such that the length of the ring is given by $L = N\epsilon$. The oscillator picture is then obtained with $L = 1$.

We transform $\tilde{\phi}$ and $\tilde{\pi}$ to real fields by a unitary transformation

$$\begin{aligned} \tilde{\phi}_l &= \frac{1}{2} \left(1 + i \right) \phi_l + \frac{1}{2} \left(1 - i \right) \phi_{-l}, \\ \tilde{\pi}_l &= \frac{1}{2} \left(1 - i \right) \pi_l + \frac{1}{2} \left(1 + i \right) \pi_{-l}. \end{aligned} \tag{9.3}$$

Then, the Hamiltonian is diagonal in the sense that all modes are decoupled

$$H = \frac{1}{2} \sum_l \frac{\Delta k}{2\pi} \left[\pi_l^2 + \omega_l^2 \phi_l^2 \right], \tag{9.4}$$

and the frequencies read

$$\omega_l \equiv \sqrt{\frac{4}{\epsilon^2} \sin^2 \left(\frac{\Delta kl\epsilon}{2} \right) + m^2}. \tag{9.5}$$

We quantize the two fields by imposing canonical commutation relations on the hermitian field operators in momentum space

$$[\boldsymbol{\phi}_l, \boldsymbol{\pi}_{l'}] = i\frac{2\pi}{\Delta k} \delta_{ll'}, \quad [\boldsymbol{\phi}_l, \boldsymbol{\phi}_{l'}] = [\boldsymbol{\pi}_l, \boldsymbol{\pi}_{l'}] = 0. \tag{9.6}$$

To describe excitations, we introduce creation and annihilation operators

$$a_l^\dagger \equiv \frac{1}{\sqrt{2\omega_l}} \left(\omega_l \boldsymbol{\phi}_l - i\boldsymbol{\pi}_l \right), \quad a_l \equiv \frac{1}{\sqrt{2\omega_l}} \left(\omega_l \boldsymbol{\phi}_l + i\boldsymbol{\pi}_l \right). \tag{9.7}$$

From (9.6) follows then $[a_l, a_{l'}^\dagger] = \frac{2\pi}{\Delta k} \delta_{ll'}$.

It is instructive to express the quantum version of the Hamiltonian (9.4) in terms of creation and annihilation operators

$$H = \sum_l \frac{\Delta k}{2\pi} \omega_l \left(a_l^\dagger a_l + \frac{1}{2} \frac{2\pi}{\Delta k} \mathbb{1} \right). \tag{9.8}$$

The second term shows that the vacuum energy diverges in the field theory limit. We will see later that the entropy of field configurations shows a similar divergence.

The lattice theory we have set up so far has two interesting limits. First, the continuum limit $\epsilon \to 0, N \to \infty$ with $L = N\epsilon$ fixed. In this case, the spatial positions become continuous in the sense of $\phi_j \equiv \phi(j\epsilon) \to \phi(x)$ with $x \in [0, L]$ and the discrete Hamiltonian (9.1) becomes the Hamiltonian of a free massive scalar field with relativistic dispersion relation (cf. (9.5)). The periodic boundary conditions then imply that the momenta are discrete but unbounded, i.e. $l \in \mathbb{Z}$.

Second, we can take the infinite volume limit $L \to \infty$ (which corresponds to $\Delta k \to 0$) and $N \to \infty$, with $\epsilon = \frac{L}{N}$ fixed. The momenta become continuous such that $\phi_l \equiv \phi(\Delta k l) \to \phi(k)$, but are restricted to the Brillouin zone $k \in \left[-\frac{\pi}{\epsilon}, \frac{\pi}{\epsilon} \right]$. Also, the commutation relations (9.6) attain a distributional character

$$[\boldsymbol{\phi}(k), \boldsymbol{\pi}(k')] = i\delta(k - k'), \quad [\boldsymbol{\phi}(k), \boldsymbol{\phi}(k')] = [\boldsymbol{\pi}(k), \boldsymbol{\pi}(k')] = 0 \tag{9.9}$$

and have to be used with respect to an integral measure $\frac{dk}{2\pi}$.

We speak of the field theory limit when we take both limits. Then, we obtain back the standard Fourier transform

$$\phi(x) = \int \frac{dk}{2\pi} e^{ikx} \tilde{\phi}(k), \quad \pi(x) = \int \frac{dk}{2\pi} e^{-ikx} \tilde{\pi}(k) \tag{9.10}$$

and a relativistic theory on an infinite space with Hamiltonian

$$H = \frac{1}{2} \int dx \left[\pi^2(x) + (\partial_x \phi(x))^2 + m^2 \phi^2(x) \right] \tag{9.11}$$

and relativistic dispersion relation $\omega(k) = \sqrt{k^2 + m^2}$.

In the following, we will work in momentum space only. We chose the convention to label momentum space expressions by the indices l, l' and l'' in the lattice theory and by the arguments k, k' and k'' in the field theory. In most cases, the field theory limit can be taken without further ado, which applies in particular to the appearing bilinear forms.

SCHRÖDINGER FUNCTIONAL FORMALISM In analogy to QM, we introduce sets of eigenstates

$$\boldsymbol{\phi}_l \ket{\phi} = \phi_l \ket{\phi}, \quad \boldsymbol{\pi}_l \ket{\pi} = \pi_l \ket{\pi}, \tag{9.12}$$

with ϕ_l and π_l being the corresponding eigenvalues, which are the main building blocks in the Schrödinger functional formalism (see [192] for a comprehensive introduction). Without loss of generality we focus on the field ϕ in the following. In the field basis $\{\ket{\phi}\}_\phi$ the momentum operator is given by a functional derivative

$$\boldsymbol{\pi}_l = -i \frac{\delta}{\delta \phi_l}, \tag{9.13}$$

and the density operator $\boldsymbol{\rho}$ has the matrix representation [187, 192]

$$\rho[\phi_+, \phi_-] = \bra{\phi_+} \boldsymbol{\rho} \ket{\phi_-}. \tag{9.14}$$

In analogy to the measured distribution $f(x)$ in QM, we define the functional probability density as the diagonal elements of the latter expression

$$F[\phi] = \rho[\phi, \phi] = \bra{\phi} \boldsymbol{\rho} \ket{\phi}. \tag{9.15}$$

One may think of $F[\phi]$ as the probability density to measure the quantum field $\boldsymbol{\phi}_l$ in the configuration ϕ_l.

As the state $\boldsymbol{\rho}$ is by definition a non-negative and normalized operator, the functional probability distribution is non-negative $F[\phi] \geq 0$ and normalized $\int \mathcal{D}\phi \, F[\phi] = \text{Tr}\{\boldsymbol{\rho}\} = 1$ as well. Note that for a pure state $\boldsymbol{\rho} = \ket{\boldsymbol{\psi}}\bra{\boldsymbol{\psi}}$ the matrix representation $\rho[\phi_+, \phi_-]$ becomes a scalar product of Schrödinger wave functionals $\Psi[\phi] = \braket{\phi | \psi}$. In this case, Born's rule applies [286]

$$F[\phi] = |\Psi[\phi]|^2, \tag{9.16}$$

which justifies the probabilistic interpretation of $F[\phi]$.

One of the advantages of the Schrödinger functional formalism is that expectation values can be computed from functional integrals

$$\langle \phi_l \rangle = \int \mathcal{D}\phi \, \phi_l \, F[\phi],$$
$$\langle \pi_l \rangle = -i \int \mathcal{D}\phi \, \frac{\delta \rho[\phi_+, \phi_-]}{\delta \phi_{+l}} \Big|_{\phi_+ = \phi_- = \phi}, \tag{9.17}$$

where we used the functional integral measure

$$\int \mathcal{D}\phi = \prod_l \int \mathrm{d}\phi_l \, \sqrt{\frac{\Delta k}{2\pi}}. \tag{9.18}$$

In the same manner one can compute n-point correlation functions. In a free theory, propagators are of special interest, which is why we label the connected two-point correlators by

$$\mathcal{M}_{ll'} = \langle \phi_l \phi_{l'} \rangle - \langle \phi_l \rangle \langle \phi_{l'} \rangle, \quad \mathcal{N}_{ll'} = \langle \pi_l \pi_{l'} \rangle - \langle \pi_l \rangle \langle \pi_{l'} \rangle. \tag{9.19}$$

VACUUM AND COHERENT STATES   We continue with setting up the functional probability densities of coherent states, which will serve as reference distribution in our field-theoretic REUR, and we start with the vacuum state $|0\rangle$. Its wave functional $\bar{\Psi}[\phi]$ is the solution of the stationary Schrödinger equation $H\bar{\Psi}[\phi] = \bar{E}\bar{\Psi}[\phi]$. Plugging in the operator-valued version of the Hamiltonian (9.4) and employing the functional derivative representation of the momentum operator (9.13) gives the expressions

$$\bar{\Psi}[\phi] = \frac{1}{\sqrt{\bar{Z}_\phi}}\exp\left(-\frac{1}{4}\sum_l\frac{\Delta k}{2\pi}\sum_{l'}\frac{\Delta k}{2\pi}\phi_l\,\bar{\mathcal{M}}_{ll'}^{-1}\,\phi_{l'}\right),$$
$$\bar{\Psi}[\pi] = \frac{1}{\sqrt{\bar{Z}_\pi}}\exp\left(-\frac{1}{4}\sum_l\frac{\Delta k}{2\pi}\sum_{l'}\frac{\Delta k}{2\pi}\pi_l\,\bar{\mathcal{N}}_{ll'}^{-1}\,\pi_{l'}\right). \tag{9.20}$$

The normalization constants evaluate to products over all single vacuum contributions

$$\bar{Z}_\phi = \prod_l\sqrt{\frac{\pi}{\omega_l}}, \quad \bar{Z}_\pi = \prod_l\sqrt{\pi\omega_l}, \tag{9.21}$$

and the inverse covariance matrices read

$$\bar{\mathcal{M}}_{ll'}^{-1} = \frac{2\pi}{\Delta k}2\omega_l\delta_{ll'}, \quad \bar{\mathcal{N}}_{ll'}^{-1} = \frac{2\pi}{\Delta k}\frac{2}{\omega_l}\delta_{ll'} \tag{9.22}$$

in the lattice theory. Note that matrices in momentum space are inverted according to

$$\sum_{l'}\frac{\Delta k}{2\pi}\bar{\mathcal{M}}_{ll'}^{-1}\bar{\mathcal{M}}_{l'l''} = \frac{2\pi}{\Delta k}\delta_{ll''}, \tag{9.23}$$

such that the covariance matrices themselves are given by

$$\bar{\mathcal{M}}_{ll'} = \frac{2\pi}{\Delta k}\frac{1}{2\omega_l}\delta_{ll'}, \quad \bar{\mathcal{N}}_{ll'} = \frac{2\pi}{\Delta k}\frac{\omega_l}{2}\delta_{ll'}. \tag{9.24}$$

One can convince oneself that the latter expressions have distributional analogs in the field theory limit, i.e.

$$\int\frac{\mathrm{d}k'}{2\pi}\,\bar{\mathcal{M}}^{-1}(k,k')\bar{\mathcal{M}}(k',k'') = 2\pi\delta(k-k'') \tag{9.25}$$

and

$$\bar{\mathcal{M}}(k,k') = \frac{\pi}{\omega(k)}\delta(k-k'), \quad \bar{\mathcal{N}}(k,k') = \pi\omega(k)\delta(k-k'). \tag{9.26}$$

As the diagonal elements of the continuous covariance matrices are formally infinite $\bar{\mathcal{M}}(k,k) \sim \delta(0)$, they have to be understood under an integral when considering the field theory limit. Dealing with these divergences is one of the main challenges for formulating an EUR for quantum fields.

The Schrödinger wave functionals $\bar{\Psi}[\phi]$ and $\bar{\Psi}[\pi]$ do only differ in their covariance matrices. Hence, we leave out the expressions for the momentum field π in the following as they can be obtained by replacing $\bar{\mathcal{M}}$ with $\bar{\mathcal{N}}$. We apply Born's rule (9.16) to compute the vacuum functional probability density

$$\bar{F}[\phi] = \frac{1}{\bar{Z}_\phi} \exp\left(-\frac{1}{2} \sum_l \frac{\Delta k}{2\pi} \sum_{l'} \frac{\Delta k}{2\pi} \phi_l \, \bar{\mathcal{M}}_{ll'}^{-1} \phi_{l'} \right). \tag{9.27}$$

Displacing the vacuum state in phase space, i.e. changing the field expectation values, generates the set of coherent states (cf. Section 3.4) and hence their functional probability densities read

$$F_\alpha[\phi] = \frac{1}{\bar{Z}_\phi} \exp\left(-\frac{1}{2} \sum_l \frac{\Delta k}{2\pi} \sum_{l'} \frac{\Delta k}{2\pi} (\phi_l - \phi_l^\alpha) \bar{\mathcal{M}}_{ll'}^{-1} (\phi_{l'} - \phi_{l'}^\alpha) \right), \tag{9.28}$$

with the notation $\langle \phi_l \rangle_\alpha \equiv \phi_l^\alpha$ and $\langle \pi_l \rangle_\alpha \equiv \pi_l^\alpha$ understood. The index α is the complex phase field

$$\alpha_l = \frac{1}{\sqrt{2}} \left(\phi_l^\alpha + i\pi_l^\alpha \right), \tag{9.29}$$

which parameterizes the set of coherent states in analogy to (3.46). Note that the vacuum is the special case $\phi_l^\alpha = \pi_l^\alpha = 0$.

9.2 RELATIVE ENTROPIC UNCERTAINTY RELATION – PART II

DIVERGENCE OF THE FUNCTIONAL ENTROPY To the functional probability density $F[\phi]$ we associate a functional entropy by summing over all field configurations in a continuous sense

$$S[F] = -\int \mathcal{D}\phi \, F[\phi] \ln F[\phi]. \tag{9.30}$$

In the field theory limit, the functional entropy $S[F]$ corresponds to the infinite-dimensional generalization of the differential entropy $S(f)$ of a single oscillator and reduces to the latter when considering $d = 0 + 1$ spacetime dimensions.

In the lattice theory with finite $N < \infty$, the functional entropy $S[F]$ is well-defined as it reduces to a multi-dimensional differential entropy. In this case, the BBM EUR can be generalized to [29]

$$S[F] + S[G] \geq N(1 + \ln \pi), \tag{9.31}$$

with a bound which scales with the total number of modes N. This shows that the continuum as well as the infinite volume limits, which both require $N \to \infty$, lead to a divergent bound, indicating that also the functional entropies are divergent.

Indeed, we find that the functional entropy $S[F]$ is divergent in the field theory limit for all quantum states. For example, when considering the vacuum state and the corresponding functional probability density (9.27), we find formally

$$
\begin{aligned}
S[\bar{F}] &= \ln \bar{Z}_\phi + \frac{1}{2}\mathrm{Tr}\left\{\bar{\mathcal{M}}^{-1}\bar{\mathcal{M}}\right\} \\
&= \ln \bar{Z}_\phi + \frac{1}{2}\int \mathrm{d}k\,\delta(0) \\
&\to \infty.
\end{aligned}
\tag{9.32}
$$

However, divergences are typical in QFTs and the present divergence is similar to the one of the vacuum energy expectation value $\bar{E} = \mathrm{Tr}\{\bar{\rho}H\}$. Using the field-theoretic Hamiltonian (9.11) we obtain

$$
\bar{E} = \frac{1}{2}\int \mathrm{d}k\,\omega(k)\delta(0) \to \infty.
\tag{9.33}
$$

Physically, this divergence arises from adding up infinitely many equal and finite contributions of oscillators in their ground states. Therefore, one argues that energies have to be measured *with respect to* the (divergent) vacuum energy, such that excitations have a finite energy by definition. In this argument, taking the vacuum energy as a reference energy appears to be the natural choice, as the vacuum state minimizes the energy.

The same logic applies to functional entropies, whose absolute values do not have any physical meaning. Instead, one has to work with relative measures when asking questions about entropic uncertainty in field theories, leading us to the notion of functional relative entropies. Then, the remaining task is to find meaningful references distributions.

FUNCTIONAL RELATIVE ENTROPY The universality of the relative entropy dictates the definition of the functional relative entropy. We consider a functional distribution $F[\phi]$ and some reference distribution $\tilde{F}[\phi]$ such that it reads

The functional relative entropy is the extension of the differential relative entropy to infinitely many degrees of freedom.

$$
S[F\|\tilde{F}] = \int \mathcal{D}\phi\, F[\phi]\left(\ln F[\phi] - \ln \tilde{F}[\phi]\right).
\tag{9.34}
$$

By an infinite-dimensional analog of Jensen's inequality, $S[F\|\tilde{F}]$ is non-negative and zero if and only if $F[\phi] = \tilde{F}[\phi]$ everywhere. As usual, we require the support condition $\mathrm{supp}\,[F(\phi)] \subseteq \mathrm{supp}\,[\tilde{F}(\phi)]$ to hold, otherwise we set $S[F\|\tilde{F}] = +\infty$. Note that this condition is always fulfilled if the reference distribution $\tilde{F}[\phi]$ is of Gaussian form, which will indeed be the case in our field-theoretic REUR.

In contrast to the functional entropy $S[F]$, the functional relative entropy $S[F\|\tilde{F}]$ is well-defined and its properties are preserved in the field theory limit, just as it was the case for the limit from discrete to continuous variables (cf. Section 2.2). Although (9.34) suggests that a definition is directly possible in the continuum theory (i.e.

without introducing an UV-regulator in the first place), this should be investigated within algebraic QFT in more detail in the future.

DERIVING THE FIELD–THEORETIC REUR Let us start with the multi-dimensional BBM relation (9.31). One of the main takeaways of Chapter 8 was that an EUR can be formulated as a REUR by choosing reference distributions $\tilde{F}[\phi]$ which maximize an entropy for a given side constraint. As we deal with a free theory, coherent states have a Gaussian functional probability density $F_\alpha[\phi]$ and hence maximize the functional entropy $S[\tilde{F}]$ for given vacuum two-point correlation functions $\bar{M}$ and field expectation values ϕ_l^α. Additionally, coherent distributions appear as natural choices as coherent states minimize all uncertainty relations. In this sense, entropic uncertainty is measured with respect to minimum uncertainty distributions just as energy is measured with respect to the vacuum.

By using $S[F_\alpha] = S[\tilde{F}]$ and (9.32) we find for the entropic uncertainty of some generic distribution $F[\phi]$ with two-point function $\mathcal{M}$ and field expectation value φ_l with respect to a coherent reference distribution a linear decomposition

$$
\begin{aligned}
S[F\|F_\alpha] &= -S[F] + \ln \bar{Z}_\phi + \frac{1}{2}\mathrm{Tr}\left\{\bar{\mathcal{M}}^{-1}\mathcal{M}\right\} \\
&= -S[F] + S[\bar{F}] \\
&\quad + \frac{1}{2}\mathrm{Tr}\left\{\bar{\mathcal{M}}^{-1}\left(\mathcal{M} - \bar{\mathcal{M}}\right)\right\} + \frac{1}{2}s\bar{\mathcal{M}}^{-1}s,
\end{aligned}
$$
(9.35)

with $s_l = \phi_l - \phi_l^\alpha$ denoting the difference of field expectation values.

It is left to decide which of the coherent distributions we should pick. As the family of coherent states is parameterized by the two field expectation values ϕ_l^α and π_l^α, we choose the coherent state such that both agree with the field expectation values of the true state ρ, which uniquely determines both reference distributions $\tilde{F}[\phi] = F_\alpha[\phi]$ and $\tilde{G}[\pi] = G_\alpha[\pi]$. These distributions minimize the functional relative entropy (9.35) with respect to s and can hence be considered as *optimal coherent references*.

With this choice, we take (9.31), plug in (9.35) for optimal coherent reference distributions and use that $S[\bar{F}] + S[\bar{G}] = N(1 + \ln \pi)$, to end up with the field-theoretic REUR

$$
\begin{aligned}
&S[F\|F_\alpha] + S[G\|G_\alpha] \\
&\leq \frac{1}{2}\mathrm{Tr}\left\{\bar{\mathcal{M}}^{-1}(\mathcal{M} - \bar{\mathcal{M}}) + \bar{\mathcal{N}}^{-1}(\mathcal{N} - \bar{\mathcal{N}})\right\}.
\end{aligned}
$$
(9.36)

The field-theoretic REUR describes many modes and fields simultaneously.

If one would have chosen coherent, but not optimal coherent references, two bilinear forms as in the last line of (9.35) would have appeared in the bound. In any case, the two distributions $F[\phi]$ and $G[\pi]$ correspond to an arbitrary state ρ.

DISCUSSION   A few comments on the field-theoretic REUR are in order.

a) *Bound does not depend on number of modes.*

The most important observation is that the bound in (9.36) does only contain differences between the true and the vacuum co-variance matrices. It does not depend on the total number of oscillator modes N and holds especially in the field theory limit. Hence, the field-theoretic REUR describes entropic uncertainty for many oscillators as well as for quantum fields.

b) *Sum of relative entropies is bounded from above.*

As we have seen already for the REUR for finite quantum systems (8.34), a sum of relative entropies has an upper bound which expresses the uncertainty principle. More precisely, the (joint) distinguishability with respect to minimum uncertainty distributions (which are at the same time maximum entropy distributions) is limited.

Also, the bound carries some state-dependence through the true two-point correlators $\mathcal{M}$ and $\mathcal{N}$. Note that this state-dependence does not mean that the bound is more tight compared to the (multi-dimensional) BBM relation, it rather expresses the fact that the reference distributions are of Gaussian form. Hence, the relation becomes tight for all (squeezed) coherent states, in which case both sides of the inequality are (finite) zero. The relevance of coherent states for uncertainty relations is emphasized in this formulation.

Furthermore, one may appreciate that the bound is of such simple form and can be computed even if the relative entropies can not. This may allow for applications which go beyond our scope.

c) *Bound is of quantum origin.*

It is instructive to consider the classical limit for the bound. To that end, we reinsert Planck's reduced constant $\hbar$ and take $\hbar \to 0$ afterwards. To shape the intuition, we start with the multi-dimensional BBM EUR (9.31) under the assumption that $N < \infty$. One finds that the lower bound disappears

$$S[F] + S[G] \geq \frac{1}{2} \ln \det \left(\hbar \mathcal{M} \cdot \hbar \mathcal{N} \right) \to -\infty, \tag{9.37}$$

as both distributions $F[\phi]$ and $G[\pi]$ can be localized arbitrarily at the same time. We refer to [28, 113] for the mindful reader who noticed that we have put an $\hbar$ inside a logarithm.

For our field-theoretic REUR (9.36), similar considerations show that in the same manner the upper bound is released in the classical limit

$$S[F\|F_\alpha] + S[G\|G_\alpha]$$

$$\leq \frac{1}{2}\mathrm{Tr}\left\{\frac{\bar{\mathcal{M}}^{-1}}{\hbar}\left(\mathcal{M} - \hbar\bar{\mathcal{M}}\right) + \frac{\bar{\mathcal{N}}^{-1}}{\hbar}\left(\mathcal{N} - \hbar\bar{\mathcal{N}}\right)\right\} \qquad (9.38)$$

$$\to +\infty.$$

Therefore, we can conclude that the bound is indeed of quantum origin.

9.3 EXAMPLES AND MEASURABILITY

EXCITED STATES   With the REUR (9.36) at hand, let us exemplify the finiteness of the bound and quantify the entropic uncertainty for some interesting states. We start with excited states. In free theories, these states often have a (quasi)-particle interpretation. We consider the most general case where we can have n_k excitations in mode k (in the field theory limit, this corresponds to n_k excitations with momentum k), where k is drawn from an index set $k \in \Im$ containing all excited modes [192]. Such a state is constructed by applying the creation operator $a_k^\dagger$ for the mode k to the vacuum n_k times for every $k \in \Im$

$$\Psi_{n_k}[\phi] = \prod_{k\in\Im} \frac{1}{\sqrt{n_k!}}\left(\sqrt{\frac{\Delta k}{2\pi}} a_k^\dagger\right)^{n_k} \Psi[\phi]. \qquad (9.39)$$

Note that all appearing factors are a matter of normalization. Using the Schrödinger representation of the momentum field operator (9.13), the creation and annihilation operators can be written as

$$a_k^\dagger = \frac{1}{\sqrt{2\omega_k}}\left(\omega_k\phi_k - \frac{\delta}{\delta\phi_k}\right), \quad a_k = \frac{1}{\sqrt{2\omega_k}}\left(\omega_k\phi_k + \frac{\delta}{\delta\phi_k}\right), \quad (9.40)$$

leading to

$$F_{n_k}[\phi] = \prod_{k\in\Im} \frac{1}{n_k!} H_{n_k}^2\left(\frac{\phi_k}{\sqrt{\bar{\mathcal{M}}_{kk}}}\right) \bar{F}[\phi], \qquad (9.41)$$

with $H_{n_k}(\phi_k)$ denoting the probabilist's Hermite polynomials

$$H_{n_k}(\phi_k) = n_k! \sum_{\gamma=0}^{\lfloor\frac{n_k}{2}\rfloor} \frac{(-1)^\gamma \phi_k^{n_k-2\gamma}}{\gamma!\,(n_k - 2\gamma)!\,2^\gamma}. \qquad (9.42)$$

It is important to note that these states are not normalizable in the field theory limit. More precisely, their overlap contains infinite factors additional to those of the vacuum state, which come again from

$\bar{\mathcal{M}}(k, k) \sim \delta(0)$. Physically, this issue is related to the fact that a free particle in a relativistic QFT has a definite momentum and is completely delocalized. We will also discuss mathematically well-defined notions of excitations later by considering averaged fields.

Let us now analyze the REUR (9.36) for excited states (9.41). We first note that $F[\phi]$ contains only even powers of ϕ, such that the Gaussian integral for the field expectation value evaluates to zero. Hence, the optimal coherent reference distribution is given by the vacuum distribution $\bar{F}[\phi]$. To compute the bound, we start from the definition of the connected two-point correlator

$$\mathcal{M}_{ll'}^{n_k} = \int \mathcal{D}\phi \left[\phi_l \phi_{l'} \prod_{k \in \mathfrak{I}} \frac{1}{n_k!} H_{n_k}^2 \left(\frac{\phi_k}{\sqrt{\bar{\mathcal{M}}_{kk}}} \right) \right] \bar{F}[\phi] \tag{9.43}$$

and consider off-diagonal and diagonal components separately. In the former case, we again deal with Gaussian integrals of odd polynomials, such that $\mathcal{M}^{n_k}$ has to be of diagonal form. To obtain the diagonal elements $\mathcal{M}_{ll}^{n_k}$, we first note that an entry corresponding to a non-excited mode ϕ_l with $l \notin \mathfrak{I}$ is given by the vacuum entry $\bar{\mathcal{M}}_{ll}$.

The only non-trivial contribution comes with excited modes ϕ_l with $l \in \mathfrak{I}$. The corresponding entry $\mathcal{M}_{ll}^{n_k}$ can be calculated by noting that $\phi_k^2 = H_1^2(\phi_k)$, then using the recurrence relation

$$H_{a+1}(\phi_l) = H_1(\phi_l) H_a(\phi_l) - a H_{a-1}(\phi_l) \tag{9.44}$$

and finally employing orthogonality between two Hermite polynomials

$$\int_{-\infty}^{+\infty} \mathrm{d}\phi_l \, H_a(\phi_l) H_b(\phi_l) e^{-\frac{1}{2}\phi_l^2} = \sqrt{2\pi} a! \delta_{ab}, \tag{9.45}$$

to obtain

$$\mathcal{M}_{ll}^{n_k} = \bar{\mathcal{M}}_{ll}(1 + 2n_l) \tag{9.46}$$

for $l \in \mathfrak{I}$.

In addition, the full two-point correlator reads

$$\mathcal{M}_{ll'}^{n_k} = \bar{\mathcal{M}}_{ll'} + \sum_{k \in \mathfrak{I}} \frac{2n_k}{\bar{\mathcal{M}}_{kk}} \bar{\mathcal{M}}_{lk} \bar{\mathcal{M}}_{l'k}, \tag{9.47}$$

showing that only components corresponding to excited modes are modified by a positive term accounting for the excitations.

We are now ready to compute the first term of the bound of the field-theoretic REUR (9.36). For the lattice theory we find

$$\begin{aligned} &\frac{1}{2}\mathrm{Tr}\left\{ \bar{\mathcal{M}}^{-1}(\mathcal{M}^{n_k} - \bar{\mathcal{M}}) \right\} \\ &= \sum_l \frac{\Delta k}{2\pi} \sum_{l'} \frac{\Delta k}{2\pi} \bar{\mathcal{M}}_{ll'}^{-1} \sum_{k \in \mathfrak{I}} \frac{n_k}{\bar{\mathcal{M}}_{kk}} \bar{\mathcal{M}}_{lk} \bar{\mathcal{M}}_{l'k} \\ &= \sum_{k \in \mathfrak{I}} \frac{n_k}{\bar{\mathcal{M}}_{kk}} \bar{\mathcal{M}}_{kk} \\ &= \sum_{k \in \mathfrak{I}} n_k, \end{aligned} \tag{9.48}$$

where we used (9.23) for inverse matrices.

In the field theory limit the correlators $\mathcal{M}^{n_k}(k', k'')$ have to be understood in a distributional sense, rendering their diagonal entries formally infinite $\mathcal{M}^{n_k}(k', k') \sim \delta(0)$. Nevertheless, the bound term remains finite as

$$
\begin{aligned}
&\frac{1}{2}\mathrm{Tr}\left\{\bar{\mathcal{M}}^{-1}(\mathcal{M}^{n_k} - \bar{\mathcal{M}})\right\} \\
&= \int \frac{\mathrm{d}k'}{2\pi} \frac{\mathrm{d}k''}{2\pi} \bar{\mathcal{M}}^{-1}(k', k'') \sum_{k \in \mathfrak{I}} \frac{n_k}{\bar{\mathcal{M}}(k,k)} \bar{\mathcal{M}}(k', k) \bar{\mathcal{M}}(k'', k) \\
&= \sum_{k \in \mathfrak{I}} \frac{n_k}{\bar{\mathcal{M}}(k,k)} \bar{\mathcal{M}}(k,k) \\
&= \sum_{k \in \mathfrak{I}} n_k.
\end{aligned}
\tag{9.49}
$$

Furthermore, we obtain exactly the same result as in the lattice theory, showing again the universality of uncertainty relations formulated in terms of relative entropy.

The contribution from the two-point correlator $\mathcal{N}^{n_k}$ of the momentum field π follows analogously as the functional probability distributions are of the same form. Thus, we end up with the simple expression for the bound

All divergences in the bound cancel out.

$$
\frac{1}{2}\mathrm{Tr}\left\{\bar{\mathcal{M}}^{-1}(\mathcal{M}^{n_k} - \bar{\mathcal{M}}) + \bar{\mathcal{N}}^{-1}(\mathcal{N}^{n_k} - \bar{\mathcal{N}})\right\} = \sum_{k \in \mathfrak{I}} 2n_k,
\tag{9.50}
$$

which does indeed not depend on the total number of modes N. For a finite number of excitations, this bound remains finite and constraints the entropic uncertainty for (quasi-)particle excitations especially of a quantum field.

Let us also analyze the functional relative entropies on the left hand side of the field-theoretic REUR (9.36). Unfortunately, the computation of an entropy or relative entropy for arbitrary excitations is very complex due to the appearance of Hermite polynomials inside the logarithm (see Section 10.2 for more details in the case of one oscillator mode). Hence, we only consider one excited mode k being excited once, which corresponds to a free particle with energy $\omega(k)$ in the continuum theory. In this case, the REUR evaluates to

$$
S[F_{1_k} \| \bar{F}] + S[G_{1_k} \| \bar{G}] = 4 - \ln 4 - 2\gamma \leq 2,
\tag{9.51}
$$

where $\gamma \approx 0.577$ denotes the Euler-Mascheroni constant. The difference between the two sides of the inequality, which we call the uncertainty deficit, agrees with the result of a single oscillator (cf. corresponding values for BBM relation in Section 10.2). A similar result is expected to hold for arbitrary excitations, but this (rather technical) discussion is omitted for brevity.

The entropic uncertainty of a single excited mode and a free particle in a relativistic field theory agree.

THERMAL STATE We also consider the thermal state (cf. (6.20)), which will add a caveat in the infinite volume limit. Its functional probability density is of Gaussian form and reads

$$F_c[\phi] = \frac{1}{Z_\phi^c} \exp\left(-\frac{1}{2} \sum_l \frac{\Delta k}{2\pi} \sum_{l'} \frac{\Delta k}{2\pi} \phi_l (\mathcal{M}_{ll'}^c)^{-1} \phi_{l'} \right), \qquad (9.52)$$

with the thermal covariance matrix

$$\mathcal{M}_{ll'}^c = (1 + 2n_{BE}(\omega_l)) \, \bar{\mathcal{M}}_{ll'}. \qquad (9.53)$$

The modes are occupied following the Bose-Einstein distribution

$$n_{BE}(\omega_l) = \frac{1}{e^{\beta \omega_l} - 1} \qquad (9.54)$$

with $n_{BE}(\omega_l) \geq 0$. In the zero temperature limit $\beta \to \infty$ we get back the vacuum expression $\lim_{T \to 0} F_c[\phi] = \bar{F}[\phi]$ as $n_{BE}(\omega_l) \to 0$ for all $\omega_l > 0$.

Starting with the lattice theory and plugging (9.53) into the bound of the REUR (9.36) yields

$$\frac{1}{2} \text{Tr} \left\{ \bar{\mathcal{M}}^{-1}(\mathcal{M}^c - \bar{\mathcal{M}}) + \bar{\mathcal{N}}^{-1}(\mathcal{N}^c - \bar{\mathcal{N}}) \right\} = 2L \sum_l \frac{\Delta k}{2\pi} n_{BE}(\omega_l). \quad (9.55)$$

Let us discuss the continuum and infinite volume limits separately. In the continuum limit $\epsilon \to 0, N \to \infty$ with $L = N\epsilon$ fixed the range of the sum over l becomes unbounded. However, the Bose-Einstein distribution $n_{BE}(\omega_l)$ decreases exponentially for large energies ω_l, for which we have $\omega_l \to l$. Hence, the series in (9.55) still converges and we obtain a finite result.

In contrast, the infinite volume limit $L \to \infty, N \to \infty$ with ϵ fixed leads to divergences outside of the integral as the discrete bound (9.55) is proportional to L. Hence, we obtain

$$\frac{1}{2} \text{Tr} \left\{ \bar{\mathcal{M}}^{-1}(\mathcal{M}^c - \bar{\mathcal{M}}) + \bar{\mathcal{N}}^{-1}(\mathcal{N}^c - \bar{\mathcal{N}}) \right\}$$
$$= 4\pi \delta(0) \int \frac{dk}{2\pi} n_{BE}(\omega(k)), \qquad (9.56)$$

with $L = \frac{2\pi}{\Delta k} \to 2\pi\delta(0)$ being an infinite volume factor. However, the integral is well-defined and finite as a result of the exponential fall-off of $n_{BE}(\omega(k))$ for large k.

Again, we note the similarities between entropic uncertainty and the energy expectation value. The latter has an infinite energy difference to the vacuum state $E_c - \bar{E} \sim \delta(0)$, which is why one either considers energy densities or puts the theory in a box, i.e. does not take the infinite volume limit.

With the same argument one can regularize the field-theoretic REUR (9.36). Hence, we can conclude that if the state of interest has finite

energy, the bound is also finite. If this is not the case, we can consider functional relative entropy densities, in which case both sides of the REUR are divided by the same infinite volume factor $\delta(0)$ to end up with finite quantities again.

At last, let us also compute the functional relative entropies. As the distributions are of Gaussian form, this is easily done. For the lattice theory we find

$$
\begin{aligned}
&S[F_c\|\bar{F}] + S[G_c\|\bar{G}] \\
&= L \sum_l \frac{\Delta k}{2\pi} \left[2n_{\mathrm{BE}}(\omega_l) - \ln\left(1 + 2n_{\mathrm{BE}}(\omega_l)\right) \right].
\end{aligned}
\tag{9.57}
$$

We plot both sides of the REUR for a single oscillator mode, which is achieved by setting $N = L = \epsilon = 1, \Delta k = 2\pi$ in Figure 9.1. The uncertainty deficit vanishes in the zero temperature limit $\beta \to \infty$ and grows boundless in the infinite temperature limit $\beta \to 0$.
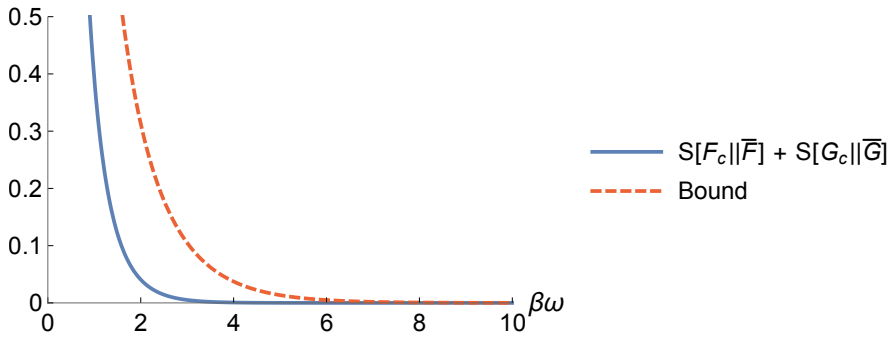
Figure 9.1: Both sides of the field-theoretic REUR (9.36) for a single mode and a thermal distribution (9.52) as a function of $\beta\omega$.

CONNECTION TO HEISENBERG'S RELATION   In Section 3.5 we have seen that the BBM EUR implies the second moment Heisenberg relation (3.80). By using relative entropy in Section 8.2 we obtained (8.28), which showed that the former is tighter than the latter whenever the measured distributions are non-Gaussian. Let us investigate how this statement can be expressed in the field-theoretic setting.

To that end, we use the same strategy as in Section 8.2 and reformulate the field-theoretic REUR by choosing Gaussian references $\tilde{F}[\phi]$ and $\tilde{G}[\pi]$ with the same expectation values and two-point correlators as the true distributions $F[\phi]$ and $G[\pi]$ instead of choosing optimal coherent reference distributions. For given correlators and expectation values, these distributions have maximal functional entropy and hence we write $\tilde{F}[\phi] = F_{\max}[\phi]$ (analogously for the momentum field π).

Then, the REUR (9.36) becomes

$$
S[F\|F_{\max}] + S[G\|G_{\max}] \leq \Delta S[F_{\max}] + \Delta S[G_{\max}],
\tag{9.58}
$$

with

$$\Delta S[F_{\max}] = S[F_{\max}] - S[\bar{F}] = \frac{1}{2} \ln \frac{\det \mathcal{M}}{\det \bar{\mathcal{M}}}, \tag{9.59}$$

being the functional entropy difference between the Gaussian reference $F_{\max}[\phi]$ and the vacuum $\bar{F}[\phi]$. Note that this relation is the exact field-theoretic analog of the REUR (8.34) discussed in Section 8.3. Again, the bound is independent of the number of modes N and hence finite for finite energy states. Combining the latter two relations, it reads

$$\Delta S[F_{\max}] + \Delta S[G_{\max}] = \frac{1}{2} \ln \frac{\det(\mathcal{M} \cdot \mathcal{N})}{\det(\bar{\mathcal{M}} \cdot \bar{\mathcal{N}})} \tag{9.60}$$

and we find the field-theoretic analog of (8.28) to be

$$\frac{\det(\mathcal{M} \cdot \mathcal{N})}{\det(\bar{\mathcal{M}} \cdot \bar{\mathcal{N}})} \geq e^{2S[F\|F_{\max}] + 2S[G\|G_{\max}]} \geq 1. \tag{9.61}$$

Note that also this formulation of Heisenberg's relation is free of divergences in the field theory limit.

COMMENT ON MEASURABILITY   Our discussion about entropic uncertainty for quantum fields has been a rather theoretical one. We treated the quantum field operator ϕ and the conjugate momentum field operator π as normal observables and assumed that the corresponding functional probability densities $F[\phi]$ and $G[\pi]$ can be measured. However, it is well-known that ϕ and π do not constitute operators, but rather operator-valued distributions.

Physically, this problem comes from considering the fields at single points in spacetime (or in momentum space). Such fields can never be measured as this requires infinite precision and hence infinite energy. Instead, any real measurement device is limited by its accuracy, which is always finite.

To formulate a QFT in terms of well-defined field operators, one takes the operator-valued distributions and applies to them a test function, which is a standard procedure in algebraic QFT [199, 287]. The test functions need to fulfill some regularity conditions and may be chosen inspired by the structural conditions of the measurement device. This leads to an averaged or wave-packet quantum field operator, which constitutes a proper observable. See [288] for a discussion of Heisenberg's uncertainty relation in a similar setup.

A convenient choice are Gaussian test functions $\mathcal{A}(p)$, which are Schwartz functions $\mathcal{A}_k(p) \in \mathscr{S}(\mathbb{R})$. Also, Gaussians often approximate experimental procedures rather well. As we work in momentum space, we choose a Gaussian centered around a momentum k, which we write as $\mathcal{A}_k(p)$, to define an averaged field operator corresponding to the momentum k as

$$\phi(\mathcal{A}_k) = \int \frac{dp}{2\pi} \mathcal{A}_k(p) \phi(p). \tag{9.62}$$

As the quantum field itself, this operator fulfills an eigenvalue equation in analogy to (9.12)

$$\boldsymbol{\phi}\left(\mathcal{A}_k\right)|\phi\left(\mathcal{A}_k\right)\rangle = \phi\left(\mathcal{A}_k\right)|\phi\left(\mathcal{A}_k\right)\rangle, \tag{9.63}$$

with $\phi\left(\mathcal{A}_k\right)$ being the classical eigenvalues.

We wish to investigate the field-theoretic REUR (9.36) for averaged quantum fields and exemplarily consider a single excitation. The corresponding averaged Schrödinger wave functional reads

$$\Psi_{1_k}[\phi] = \frac{\phi[\mathcal{A}_k]}{\sqrt{\bar{\mathcal{M}}(k,k)}}\bar{\Psi}[\phi] \tag{9.64}$$

and by Born's rule (9.16) we can compute the functional probability density

$$F_{1_k}[\phi] = \frac{\phi^2[\mathcal{A}_k]}{\bar{\mathcal{M}}(k,k)}\bar{F}[\phi], \tag{9.65}$$

which is a measurable object, in principle. Its normalization constant Z_ϕ^1 fulfills

$$Z_\phi^1 = \bar{Z}_\phi\,\bar{\mathcal{M}}(k,k), \tag{9.66}$$

but the diagonal entries of the vacuum two-point correlator are now finite as

$$\bar{\mathcal{M}}(k,k) = \pi \int \frac{\mathrm{d}k'}{2\pi}\frac{\mathcal{A}_k^2(k')}{\omega(k')} < \infty. \tag{9.67}$$

As a consequence, not only the bound of the REUR is finite, but also all terms appearing during the calculation. Omitting a few technical steps, we find that

$$\begin{aligned}
\frac{1}{2}\mathrm{Tr}\left\{\bar{\mathcal{M}}^{-1}\left(\mathcal{M}^{1_k} - \bar{\mathcal{M}}\right)\right\} &= \frac{\pi}{\bar{\mathcal{M}}(k,k)}\int\frac{\mathrm{d}k'}{2\pi}\frac{\mathcal{A}_k^2(k')}{\omega(k')} \\
&= 1,
\end{aligned} \tag{9.68}$$

where we used (9.67) in the second step.

Hence, we can conclude that the REUR (9.36) does not only accurately describe entropic uncertainty for many modes and quantum fields, but also for averaged quantum fields. Actually, the bound attains the same value in all cases, emphasizing again the universality of formulating uncertainty relations with relative entropy.

10 | ENTROPIC UNCERTAINTY IN PHASE SPACE

This chapter is taken from the first half of [D]. I proposed the project and S. F. served as supervisor. The calculations were done by H. M.-G. and me and I produced all figures. All authors contributed to the writing, while I wrote large parts of the text. Most of what is presented here is also covered in the bachelor thesis by H. M.-G. [289].

In the previous two chapters we have discussed entropic uncertainty in terms of distributions obtained after measuring two observables separately, i.e. following the homodyne protocol. However, the heterodyne protocol, i.e. a POVM consisting of pure coherent state projectors as introduced in Section 3.4 for position and momentum, provides an alternative. In contrast to the former approach, it outputs a full phase space distribution, namely the Husimi Q-distribution, which may be advantageous for certain questions. In particular, this distribution contains information about the correlations between the two observables, which is not included in the marginal distributions.

In the following, we ask whether the EUR in phase space (3.84) is closer to equality for typical states than the EURs by BBM (3.79) and FL (3.82) based on marginal distributions. Our discussion focuses on the continuous variables position and momentum. We leave a similar analysis for other observables for future work.

Our question is not only interesting by itself, but also paves the ground towards tighter separability criteria. In simple words, tighter uncertainty relations allow for better separability criteria in the sense that more entangled states can be detected. As the main conclusion of this chapter is the conjecture that the WL inequality is tighter than the usual BBM relation almost everywhere, we derive new and strong separability criteria in phase space in Part iv.

10.1 SETTING UP THE COMPARISON

ENTROPIES IN PHASE SPACE We consider a continuous variable quantum system with position X and momentum K. While the heterodyne detection gives us access to $Q(x, k)$, the homodyne detection gives us $f(x)$ and $g(k)$. For all distributions we know EURs, namely the WL inequality for $S(Q)$ and the BBM and FL relations for $S(f) + S(g)$.

On first sight, $S(Q)$ and $S(f) + S(g)$ are rather different quantities, as the first is an entropy of a two-dimensional phase space distribution and the latter is a sum of one-dimensional entropies. This raises the question if the comparison of these quantities is feasible at all.

However, one can translate the two marginal distributions $f(x)$ and $g(k)$ (uniquely) into the phase space distribution $f(x)g(k)$, which by definition neglects the correlations between X and K. As such, it corresponds to the Wigner W-distribution in a special case.

Then, the two-dimensional phase space entropies of the distrbutions $Q(x,k)/(2\pi)$ and $f(x)g(k)$ are given by

$$
\begin{aligned}
S\left(\frac{Q}{2\pi}\right) &= -\int \mathrm{d}x\,\mathrm{d}k\,\frac{Q(x,k)}{2\pi}\ln\left(\frac{Q(x,k)}{2\pi}\right) \\
&= S(Q) + \ln(2\pi), \\
S(fg) &= \int \mathrm{d}x\,\mathrm{d}k\,f(x)g(k)\ln\left(f(x)g(k)\right) \\
&= S(f) + S(g).
\end{aligned}
\tag{10.1}
$$

This shows that they are, up to additive constants, related with the left hand sides of the three EURs (3.84), (3.79) and (3.82). In this sense, all three EURs and in particular the two EURs (3.79) and (3.82) containing only entropies of marginal distributions make a statement about entropic uncertainty in the full phase space.

The BBM and FL relations describe entropic uncertainty in phase space for x and k being uncorrelated.

A NOTE ON TIGHTNESS We analyze the tightness of the three EURs in terms of the uncertainty deficit. Hence, we investigate how close the entropies on the left hand side are to the bounds set by the uncertainty principle. A small uncertainty deficit is favorable as this indicates the strength of the statement. Let us emphasize that this does not necessarily mean that if one relation is tighter than another one for certain states that one can deduce a formal implication between them.

Also, let us recall that the three EURs are tight, i.e. reach the bound, for different sets of states. The WL inequality is tight only for pure coherent state projectors, while the BBM relation is additionally tight for squeezed coherent states provided that the (x, k) axes are aligned with the principal axes. In contrast, the FL relation is tight only in the infinite temperature limit.

It is an interesting open problem to find an EUR which is tight for all pure Gaussian states. Note also that such a relation exists in terms of second moments (cf. (3.75)), but not in terms of entropies.

REARRANGING THE EURS As the three EURs of our interest have different right hand sides, we rearrange them accordingly. In this chapter, we compare the following three relations

$$
\left.
\begin{array}{ll}
\text{WL} & S(Q) + \ln \pi \\
\text{BBM} & S(f) + S(g) \\
\text{FL} & S(f) + S(g) - S(\rho) + \ln e/2
\end{array}
\right\} \geq \ln e\pi.
\tag{10.2}
$$

We study the relations systematically by considering (pure and mixed) number eigenstates, which form a basis and hence allow us to test several regions of the Hilbert space. For simplicity, we restrict ourselves to states which are diagonal in this basis.

10.2 PURE NUMBER EIGENSTATES

STATES AND DISTRIBUTIONS We start our analysis with pure number eigenstates

$$\rho = |n\rangle \langle n|. \tag{10.3}$$

The state vector $|n\rangle$ can be obtained from the vacuum $|0\rangle$ defined through $a|0\rangle = 0$ by repeatedly applying the creation operator $a^\dagger$

$$|n\rangle = \frac{1}{\sqrt{n!}} \left(a^\dagger\right)^n |0\rangle. \tag{10.4}$$

As the state is pure $S(\rho) = 0$, the FL EUR is always less tight than the BBM relation.

To compute the Husimi Q-distribution, it is useful to note that a pure coherent state $|\alpha\rangle$ can be expanded in terms of number eigenstates

$$|\alpha\rangle = \sum_{n=0}^{\infty} e^{-\frac{1}{2}(x^2+k^2)} \frac{(x+ik)^n}{\sqrt{2^n n!}} |n\rangle, \tag{10.5}$$

where we used the parameterization (3.46). Hence, the overlap between a coherent state and a number eigenstate is given by

$$\langle n|\alpha\rangle = e^{-\frac{1}{2}(x^2+k^2)} \frac{(x+ik)^n}{\sqrt{2^n n!}}, \tag{10.6}$$

which leads to

$$Q_n(x,k) = |\langle n|\alpha\rangle|^2 = \frac{(x^2+k^2)^n}{2^n n!} e^{-\frac{1}{2}(x^2+k^2)}. \tag{10.7}$$

The probability density functions $f(x)$ and $g(k)$ have equal shape and read

$$f_n(x) = |\langle n|x\rangle|^2 = \frac{H_n^2(x)}{\sqrt{\pi}\, 2^n n!} e^{-x^2}, \tag{10.8}$$

where $H_n(x)$ denotes the physicist's Hermite polynomials.

MARGINAL ENTROPY Computing the entropies $S(f_n)$ and $S(g_n)$ is a surprisingly hard task, which can be traced back to the appearance of the Hermite polynomials $H_n(x)$ inside the logarithm [290–293]. More precisely, one can easily confirm that the entropy evaluates to

$$S(f_n) = \ln\left(\sqrt{\pi} 2^n n!\right) + n + \frac{1}{2} + \frac{1}{\sqrt{\pi} 2^n n!} E(H_n). \tag{10.9}$$

However, it is complicated to compute $E(H_n)$, which denotes the Hermite polynomial entropy. At the end one finds

$$\begin{aligned}
E(H_n) &= -\int_{-\infty}^{+\infty} dx\, e^{-x^2} H_n^2(x) \ln H_n^2(x) \\
&= n\gamma - 2\sum_{i=1}^{n} x_{n,i}^2 \,_2F_2\left(1,1;\frac{3}{2},2;-x_{n,i}^2\right) \\
&\quad + \sum_{k=1}^{n} \binom{n}{k} \frac{(-1)^k 2^k}{k} \sum_{i=1}^{n} {}_1F_1\left(k,\frac{1}{2},-x_{n,i}^2\right),
\end{aligned} \tag{10.10}$$

where again $\gamma \approx 0.577$, $_2F_2(a, b; c, d; z)$ is a generalized hypergeometric function, $_1F_1(a, b; z)$ is Kummer's confluent hypergeometric function for $z \in \mathbb{C}$ and $x_{n,i}$ are the roots of the n-th Hermite polynomial $H_n(x^2)$. Note that this expression has to be evaluated numerically for all $n \in \mathbb{N}$.

Let us also investigate the limiting behavior for large $n \gg 1$. Although the Hermite polynomial entropy (10.10) is a complicated expression, one can find an approximation for the differential entropy

$$S(f_n) \approx \frac{1}{2} \left(-2 + \ln 2\pi^2 n \right) \tag{10.11}$$

for $n \gg 1$. See also [291] for a numerical investigation.

WEHRL ENTROPY   Interestingly, the Wehrl entropy $S(Q_n)$ can be computed analytically in a straight forward manner. Note that this is often the case and is a general advantage over the marginal entropies $S(f_n)$ and $S(g_n)$. After performing a few Gaussian integrals one ends up with (see also [294])

$$S(Q_n) = 1 + n \left(1 + \gamma - \eta_n \right) + \ln n! \tag{10.12}$$

with η_n being the nth harmonic number, i.e.

$$\eta_n = \sum_{j=1}^{n} \frac{1}{j}, \tag{10.13}$$

with the convention $\eta_0 = 0$. In the limit of large $n \gg 1$, we can apply Stirling's formula $n! \approx \sqrt{2\pi n} \left(\frac{n}{e} \right)^n$, which for $n \gg 1$ leads to the simple result

$$S(Q_n) \approx \frac{1}{2} \left(1 + \ln 2\pi n \right). \tag{10.14}$$

COMPARISON   We plot the three resulting EURs with the expressions (10.9) and (10.12) together with the bound $\ln e\pi$ in Figure 10.1a. Besides the trivial result that the FL relation is worst for pure number eigenstates, we find that the WL inequality is tighter than the BBM relation for $n > 1$. For $n = 0$, the two agree, while for $n = 1$ the BBM relation is tightest. We can conclude that the WL relation is tightest almost everywhere.

Also, we plot the WL relation against the BBM relation together with their expressions for large $n \gg 1$ in Figure 10.1b, showing that the latter grows much faster than the former. Further, we can read off that the approximation (10.14) works rather well for $n \sim 1$, while (10.11) is approached by (10.9) only for sufficiently large $n \gg 1$.

(a) Small n

(b) Large $n \gg 1$
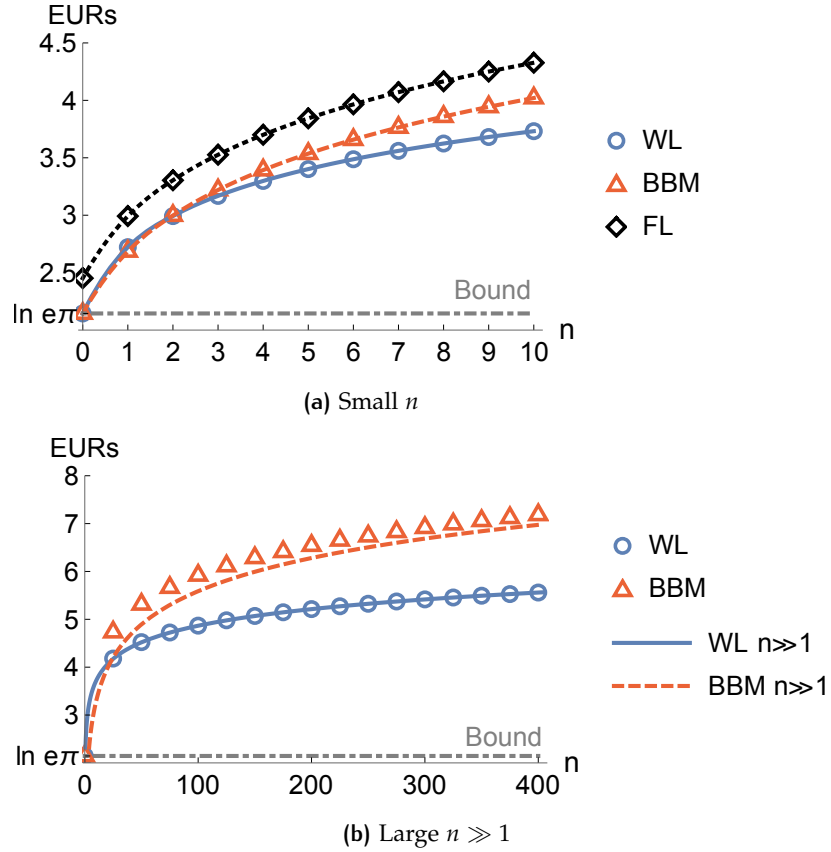
Figure 10.1: (a) EURs in (10.2) for pure number eigenstates (10.3) with $n \in \{0, ..., 10\}$. Note that for $n = 1$ we obtain $S(Q_1) + \ln \pi \approx 2.72 > S(f_1) + S(g_1) \approx 2.69$, while for $n = 2$ we have $S(Q_2) + \ln \pi \approx 2.992 < S(f_2) + S(g_2) \approx 2.997$. The curves represent interpolations. (b) WL and BBM relations with their approximations for large $n \gg 1$ (indicated by the two curves).

10.3 MIXTURES OF NUMBER EIGENSTATES

MIXING TWO NUMBER EIGENSTATES Let us now investigate mixtures of number eigenstates. We start with mixing two number eigenstates $|n\rangle$ and $|m\rangle$, such that the density operator reads

$$\rho_{nm} = q|n\rangle \langle n| + (1-q)|m\rangle \langle m|. \tag{10.15}$$

Therein, $q \in [0, 1]$ denotes the probability to obtain the pure state $|n\rangle \langle n|$ and can be interpreted as a mixing parameter. The marginal entropies $S(f_{nm})$ and $S(g_{nm})$ as well as the Wehrl entropy $S(Q_{nm})$ have to be computed numerically for fixed n, m as in both cases one ends up with a logarithm of a sum. The von Neumann entropy evaluates to a Shannon entropy $S(\rho) = S(q)$. Exemplarily, we show the three EURs for $n = 0$ and $m = 1$ in Figure 10.2a, for $n = 2$ and $m = 5$ in Figure 10.2b and for $n = 6$ and $m = 7$ in Figure 10.2c. We observe that WL inequality is closer to equality compared to the BBM relation for almost all $q \in [0, 1]$. Only for $q \approx 0$ and $m = 1$, which corresponds

(a) $n = 0$ and $m = 1$

(b) $n = 2$ and $m = 5$

(c) $n = 6$ and $m = 7$
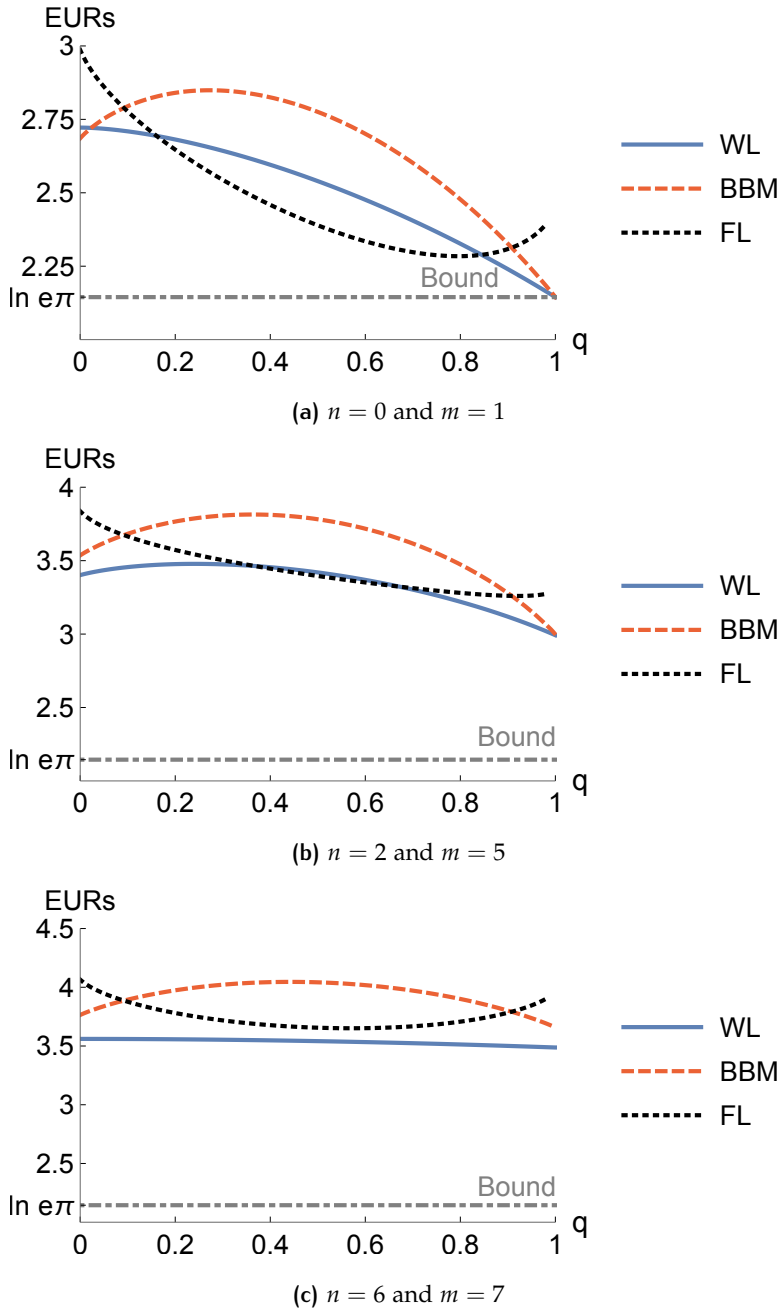
Figure 10.2: EURs in (10.2) for a mixture of two number eigenstates (10.15). The WL relation is tighter than the BBM relation almost everywhere, while the FL relation can become tightest for highly mixed and low-lying number eigenstates.

to $\rho \approx |1\rangle \langle 1|$, the BBM relation is tighter, which is accordance with our findings in Figure 10.1. We also find that the WL inequality is less concave in q compared to the BBM relation. Finally, the relation by Frank and Lieb is tightest for sufficiently mixed states, i.e. around $q \approx 0.5$, but becomes less tight than the WL relation when mixing higher excitations.

THERMAL STATE At last, we consider the thermal state for the Hamiltonian of a quantum harmonic oscillator $H = \omega \left(N + \frac{1}{2}\right)$ with $N = a^\dagger a$ being the particle number operator. It reads

$$\rho_c = \frac{1}{Z_c} e^{-\beta H} = \frac{1}{Z_c} \sum_{n=0}^{\infty} e^{-\beta \omega \left(n + \frac{1}{2}\right)} |n\rangle \langle n|, \tag{10.16}$$

with

$$Z_c = \sum_{n}^{\infty} e^{-\beta \omega \left(n + \frac{1}{2}\right)} = \frac{1}{2 \sinh \frac{\beta \omega}{2}} \tag{10.17}$$

ensuring normalization of the state.

The marginal distributions $f_c(x)$ and $g_c(p)$ are again of the same form and can be computed using Mehler's formula [295]. One finds that the distribution $f(x)$ is a centered Gaussian

$$f_c(x) = \frac{1}{Z} \sum_{n=0}^{\infty} |\langle n|x\rangle|^2 e^{-\beta \omega \left(n + \frac{1}{2}\right)} = \frac{1}{\sqrt{2\pi \sigma_c^2}} e^{-\frac{1}{2}\left(\frac{x}{\sigma_c}\right)^2}, \tag{10.18}$$

with a variance of the form

$$\sigma_c^2 = \left(2 \tanh \frac{\beta \omega}{2}\right)^{-1} \tag{10.19}$$

Then, the sum of the two marginal entropies is given by

$$S(f_c) + S(g_c) = 1 + \ln \pi - \ln \left(\tanh \frac{\beta \omega}{2}\right). \tag{10.20}$$

As expected, the BBM relation becomes tight in the zero temperature limit $\beta \to \infty$.

For large temperatures, the thermal state is highly mixed, which is expressed through the von Neumann entropy

$$S(\rho_c) = -\ln \left(1 - e^{-\beta \omega}\right) + \frac{\beta \omega}{e^{\beta \omega} - 1}, \tag{10.21}$$

such that the left hand side of the FL relation becomes

$$\begin{aligned} &S(f_c) + S(g_c) - S(\rho_c) + \ln \frac{e}{2} \\ &= 2 + \ln \left(\frac{\pi}{2} \frac{1 - e^{-\beta \omega}}{\tanh \frac{\beta \omega}{2}}\right) - \frac{\beta \omega}{e^{\beta \omega} - 1} \end{aligned} \tag{10.22}$$

The Husimi Q-distribution can be calculated in a simpler way by using the MacLaurin series of the exponential function

$$\begin{aligned} Q_c(x,k) &= \frac{1}{Z_c} \sum_{n=0}^{\infty} \frac{(x^2 + k^2)^n}{2^n n!} e^{-\frac{1}{2}(x^2 + k^2)} e^{-\beta \omega \left(n + \frac{1}{2}\right)} \\ &= \frac{1}{Z_c} e^{-\frac{1}{2}(x^2 + k^2)(1 - e^{-\beta \omega}) - \frac{\beta \omega}{2}} \end{aligned} \tag{10.23}$$

and is also of Gaussian form. We find for the left hand side of the WL inequality

$$S(Q_c) + \ln \pi = 1 + \frac{\beta\omega}{2} + \ln\left(\frac{\pi}{2}\operatorname{csch}\frac{\beta\omega}{2}\right). \tag{10.24}$$

As for the BBM relation, the WL inequality becomes tight in the zero temperature limit $\beta \to \infty$.

The three resulting EURs are shown in Figure 10.3. We find that the WL relation is closer to equality than the BBM relation and that both become tight for $\beta \to \infty$. Only for high temperatures the FL relation is tightest and becomes tight for $\beta \to 0$.
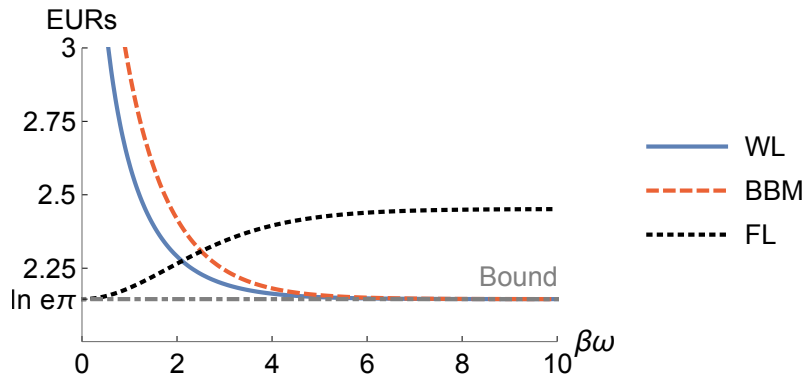
Figure 10.3: EURs in (10.2) for the thermal state (10.16). The WL relation is tighter than the BBM relation. Both become tight for $\beta \to \infty$. For sufficiently high temperatures, the FL relation becomes tightest as the state becomes more mixed. Also, it is tight in the infinite temperature limit $\beta \to 0$.

GENERAL COMMENT AND A CONJECTURE   Extrapolating the concrete evidence we have collected in the previous two sections allows us to conjecture that the WL inequality is tightest almost everywhere. To be more precise, we consider the subspaces of the full Hilbert space $\mathcal{H}', \mathcal{H}'' \subset \mathcal{H}$ for which the BBM or the FL relations are closer to equality than the WL inequality, respectively. Then, our conjecture is that $D', D'' \ll D$. We have more evidence for the first statement, which is the important one regarding entanglement witnessing. A formal proof for both statements, which is expected to be rather complicated especially in the latter case, is left for future work.

Is the WL relation tightest almost everywhere?

Also, we have seen that the Wehrl entropy is often easier to compute than marginal entropies, which might be of relevance depending on the application.

.

Part IV

ENTANGLEMENT

We derive new separability criteria in phase space in terms of the Husimi Q-distribution. For pure states, we consider the Wehrl mutual information (Chapter 11), while for mixed states we deduce entropic criteria (Chapter 12), which are then generalized to concave functions (Chapter 13)

11

ANALYSIS OF ENTROPIC WEHRL QUANTITIES

We follow the second half of [D], which is not contained in the bachelor thesis by H. M.-G. [289]. The project was supervised by S. F. and proposed by me. Most calculations were done by me, some together with H. M.-G.. I created all figures and wrote the early versions of the manuscript. All authors contributed in finalizing the draft.

We now turn to the phenomenon entanglement, which is a central feature of quantum systems. Again, we consider the continuous variables position and momentum, for which we discussed several separability criteria to detect entanglement in Section 3.6. What all reviewed approaches have in common is that they rely on the detection of the two marginal distributions $f(x)$ and $g(k)$.

In the following, we investigate how separability criteria can be formulated starting from the Husimi Q-distribution. Besides work on entanglement monotones for pure states based on a variant of the Wehrl entropy in [296], the Husimi Q-distribution has not been considered in this regard so far. We may speculate that this might be a consequence of the Husimi Q-distribution being the result of a *joint* measurement, which does not allow to violate any Bell inequality [297–299]. Nevertheless, the Husimi Q-distribution contains the full information about the underlying quantum state and hence one can indeed formulate separability criteria.

Also, in the previous chapter we have seen that the WL inequality is often tighter than the BBM relation, which is typically used to obtain entropic separability criteria (e.g. the WTSTD criteria (3.108)). As uncertainty relations are fundamentally linked to separability criteria and tighter uncertainty relations typically lead to stronger separability criteria, it is worth to study the Husimi Q-distribution in this regard.

In this chapter, we analyze the Wehrl conditional entropy and the Wehrl mutual information in Section 11.1, which turn out to be useful for the detection of pure entangled states as exemplified in Section 11.2.

11.1 CONDITIONAL ENTROPY AND MUTUAL INFOR– MATION

BIPARTITE SYSTEM FOR ARBITRARY MODES We consider the most general setup, i.e. a bipartition 12 where the subsystems 1 and 2 consist of N and M modes, respectively. Every single mode l is described

by the two degrees of freedom X_l and K_l, which fulfill the Heisenberg algebra $[X_l, K_{l'}] = i\delta_{ll'}$. Typically, these operators correspond to position and momentum. Then, we can associate vectors $\vec{X}_j = (X_1, ..., X_N)$ and $\vec{K}_j = (X_{N+1}, ..., X_{N+M})$ and Hilbert spaces $\mathcal{H}_j$ with every subsystem $j \in \{1, 2\}$. The coherent state of a subsystem is given by a tensor product over single coherent states, for example

$$|\vec{\alpha}\rangle = \bigotimes_{l=1}^{N} |\alpha_l\rangle \tag{11.1}$$

for subsystem 1. Similarly, we associate $|\vec{\beta}\rangle$ with subsystem 2, such that the global coherent states read $|\vec{\alpha}\vec{\beta}\rangle = |\vec{\alpha}\rangle \otimes |\vec{\beta}\rangle$.

The global Husimi Q-distribution is $(N + M)$-dimensional and is calculated from the global state ρ_{12}, i.e. $Q_{12}(\vec{x}_1, \vec{k}_1, \vec{x}_2, \vec{k}_2) = \langle \vec{\alpha}\vec{\beta}|\rho_{12}|\vec{\alpha}\vec{\beta}\rangle$. The local Husimi Q-distributions can either by obtained from the local states or from integrating out the complementary subsystem in the global Husimi Q-distribution. For example, for subsystem 1 we have

$$Q_1(\vec{x}_1, \vec{k}_1) = \langle \vec{\alpha}|\rho_1|\vec{\alpha}\rangle = \int \frac{d\vec{x}_2\, d\vec{k}_2}{(2\pi)^M} Q_{12}(\vec{x}_1, \vec{k}_1, \vec{x}_2, \vec{k}_2). \tag{11.2}$$

We also consider conditional Husimi Q-distributions, which describe one subsystem after having measured the other one first. If a heterodyne measurement of subsystem 2 produces the outcome $\vec{\beta}$, the conditional Husimi Q-distribution of subsystem 1 is given by

$$\begin{aligned} Q_1(\vec{x}_1, \vec{k}_1|\vec{x}_2, \vec{k}_2) &= \mathrm{Tr}\left\{ \frac{\langle\vec{\beta}|\rho_{12}|\vec{\beta}\rangle}{\langle\vec{\beta}|\rho_2|\vec{\beta}\rangle} |\vec{\alpha}\rangle \langle\vec{\alpha}| \right\} \\ &= \frac{\langle\vec{\alpha}\vec{\beta}|\rho_{12}|\vec{\alpha}\vec{\beta}\rangle}{\langle\vec{\beta}|\rho_2|\vec{\beta}\rangle} \\ &= \frac{Q_{12}(\vec{x}_1, \vec{k}_1, \vec{x}_2, \vec{k}_2)}{Q_2(\vec{x}_2, \vec{k}_2)}, \end{aligned} \tag{11.3}$$

which is in complete analogy to classical probability distributions (cf. Section 2.1). Note that $\langle\vec{\beta}|\rho_{12}|\vec{\beta}\rangle$ is an operator acting on $\mathcal{H}_1$.

WEHRL RELATIVE ENTROPY   To define the Wehrl conditional entropy and the Wehrl mutual information properly, we define the Wehrl relative entropy first. In complete analogy to its standard definition it is defined for two Husimi Q-distributions $Q_1(\vec{x}_1, \vec{k}_1)$ and $\tilde{Q}_1(\vec{x}_1, \vec{k}_1)$ (for brevity we only consider subsystem 1 here)

$$\begin{aligned} &S(Q_1\|\tilde{Q}_1) \\ &= \int \frac{d\vec{x}_1\, d\vec{k}_1}{(2\pi)^N} Q_1(\vec{x}_1, \vec{k}_1) \left(\ln Q_1(\vec{x}_1, \vec{k}_1) - \ln \tilde{Q}_1(\vec{x}_1, \vec{k}_1) \right). \end{aligned} \tag{11.4}$$

If the support condition $\mathrm{supp}\left[Q_1(\vec{x}_1, \vec{k}_1)\right] \subseteq \mathrm{supp}\left[\tilde{Q}_1(\vec{x}_1, \vec{k}_1)\right]$ is violated, we set to $S(Q_1\|\tilde{Q}_1) = +\infty$. As usual, the Wehrl relative entropy

serves as a measure for the distinguishability between the two Husimi Q-distributions. Although the Husimi Q-distribution is bounded from above, the facts that it is non-negative and normalized suffice to show that the Wehrl relative entropy is non-negative. More precisely, applying Jensen's inequality to the convex function $f(t) = -\ln t$ for $t \geq 0$ indeed gives [12]

$$
\begin{aligned}
S(Q_1 \| \tilde{Q}_1) &\geq -\ln \left(\int \frac{\mathrm{d}\vec{x}_1 \, \mathrm{d}\vec{k}_1}{(2\pi)^N} Q_1(\vec{x}_1, \vec{k}_1) \frac{\tilde{Q}_1(\vec{x}_1, \vec{k}_1)}{Q_1(\vec{x}_1, \vec{k}_1)} \right) \\
&= -\ln \left(\int \frac{\mathrm{d}\vec{x}_1 \, \mathrm{d}\vec{k}_1}{(2\pi)^N} \tilde{Q}_1(\vec{x}_1, \vec{k}_1) \right) \\
&= 0,
\end{aligned}
\tag{11.5}
$$

where we assumed the support condition to be fulfilled. As $-\ln t$ is strictly convex for $t > 0$, equality holds if and only if the two Husimi Q-distributions agree.

Let us also note that the Wehrl relative entropy does not fulfill the monotonicity property of the quantum relative entropy. In fact, not even a weaker property, namely invariance under unitary transformations, holds.

WEHRL CONDITIONAL ENTROPY   In Section 3.6 we have seen that separable states have a non-negative quantum conditional entropy, while entangled states may have have a negative quantum conditional entropy. In general, it is expected that correlations between the two subsystems manifest in reducing conditional entropies (classical and quantum).

Also, conditional classical entropies of measured distributions of non-commuting observables, are, just as classical entropies, constrained by EURs. In this case, one of the two subsystem plays the role of a memory system and one speaks of an EUR in the presence of memory. We will analyze the Wehrl conditional entropy in these two regards in the following.

As we deal with infinite-dimensional Hilbert spaces, we need to impose finiteness of all involved entropies to have well-defined quantities [75]. Assuming $S(Q_1) < \infty$ allows us to the Wehrl conditional entropy in analogy to (2.18)

$$
S(Q_1 | Q_2) = S(Q_1) - S(Q_{12} \| Q_1 \times Q_2),
\tag{11.6}
$$

which can be rewritten as

$$
S(Q_1 | Q_2) = S(Q_{12}) - S(Q_2)
\tag{11.7}
$$

provided that $S(Q_2) < \infty$. It is well-known that the Wehrl conditional entropy is non-negative, or, in other words, that the Wehrl entropy is monotonous with respect to the partial trace [23]

$$
S(Q_1 | Q_2) \geq 0,
\tag{11.8}
$$

which is a consequence of $Q_1(\vec{x}_1, \vec{k}_1) \geq Q_{12}(\vec{x}_1, \vec{k}_1, \vec{x}_2, \vec{k}_2)$.

However, we can easily show that there exists a stronger bound as a consequence of the uncertainty principle. We rewrite the Wehrl conditional entropy with the conditional Husimi Q-distribution

$$
\begin{aligned}
S(Q_1|Q_2) = - &\int \frac{\mathrm{d}\vec{x}_2\,\mathrm{d}\vec{k}_2}{(2\pi)^M} Q_2(\vec{x}_2, \vec{k}_2) \\
&\times \int \frac{\mathrm{d}\vec{x}_1\,\mathrm{d}\vec{k}_1}{(2\pi)^N} Q_1(\vec{x}_1, \vec{k}_1 | \vec{x}_2, \vec{k}_2) \ln Q_1(\vec{x}_1, \vec{k}_1 | \vec{x}_2, \vec{k}_2),
\end{aligned}
\tag{11.9}
$$

which is in complete analogy to the other standard definition (2.17). As the conditional Husimi Q-distribution is a Husimi Q-distribution itself, we can apply the WL inequality to the second row of the latter equation to obtain

$$
S(Q_1|Q_2) \geq N, \tag{11.10}
$$

The Wehrl conditional entropy is bounded from below by the uncertainty principle.

which is an EUR for the Wehrl conditional entropy in the presence of classical memory. This shows that the Wehrl conditional entropy behaves like a classical conditional entropy and that it cannot be used for entanglement witnessing in the same way as the quantum conditional entropy. Nevertheless, its behavior for correlated states is quite surprising and will be investigated in more detail in Section 11.2.

Note also that in [300] (see also [138, 139]), an EUR in the presence of quantum memory was found, for which classical-quantum states and corresponding Wehrl entropies have been used. However, such a relation can not be utilized easily in an experiment as it requires the knowledge of the quantum state of one of the two subsystems.

WEHRL MUTUAL INFORMATION   The Wehrl mutual information is defined as the Wehrl relative entropy with respect to a product Husimi Q-distribution, i.e.

$$
I(Q_1 : Q_2) = S(Q_{12} \| Q_1 \times Q_2). \tag{11.11}
$$

The Wehrl mutual information is non-negative and zero if and only if the global Husimi Q-distribution factorizes, which corresponds to the global state being a product state $\boldsymbol{\rho}_{12} = \boldsymbol{\rho}_1 \otimes \boldsymbol{\rho}_2$. One can rewrite it as

$$
\begin{aligned}
I(Q_1 : Q_2) &= S(Q_1) + S(Q_2) - S(Q_{12}) \\
&= S(Q_1) - S(Q_1|Q_2),
\end{aligned}
\tag{11.12}
$$

The Wehrl mutual information is a measurable perfect witness for pure state entanglement.

provided that all involved Wehrl entropies are finite.

The Wehrl mutual information measures quantum and classical correlations. Hence, if we restrict ourselves to globally pure states $S(\boldsymbol{\rho}_{12}) = 0$, $I(Q_1 : Q_2)$ quantifies only quantum correlations between the two subsystems being zero if and only if entanglement is not present. Hence, we can conclude that $I(Q_1 : Q_2) = 0$ is a necessary and sufficient condition for pure states to be separable.

It should be noted that for pure states many necessary and sufficient criteria exists. This includes the PPT criterion (3.98) as well as the quantum conditional entropy being non-negative (3.97). However, these criteria require the knowledge of the full quantum state ρ_{12}. The crucial advantage of our approach lies in the measurability of the Husimi Q-distribution.

Furthermore, let us remark that the Wehrl mutual information is neither an entanglement measure, nor an entanglement monotone, as it does not fulfill the minimum requirement of invariance under local unitaries

$$I(Q_1' : Q_2') \neq I(Q_1 : Q_2) \tag{11.13}$$

with $\rho_{12}' = (U_1 \otimes U_2)\rho_{12}(U_1^\dagger \otimes U_2^\dagger)$ where U_1 and U_2 are unitary operators. A counterexample is given by $U_1 = \mathbb{1}$ and $U_2 = S(a)$ with $S(a)$ being the squeezing operator on subsystem 2 and $a > 0$.

However, there is an interesting relation between the Wehrl mutual information and the quantum mutual information [127]

$$I(Q_1 : Q_2) \leq I(\rho_1 : \rho_2), \tag{11.14}$$

showing that the true quantum and classical correlations are never overestimated by the Wehrl mutual information. For pure states, we obtain the inequality

$$\frac{1}{2}I(Q_1 : Q_2) \leq S(\rho_1) = S(\rho_2). \tag{11.15}$$

The Wehrl mutual information tells how much a globally pure state is entangled at least.

Although the Wehrl mutual information is only a perfect entanglement witness and not an entanglement measure, it is a lower for the entanglement entropy. Hence, it can not only be used to decide whether a globally pure state is entangled or not, it also provides a lower bound on the entanglement measure for pure states. If the global state is not pure, (11.14) allows to make a statement about how much the state is correlated at least.

NOTE ON MARGINAL MUTUAL INFORMATIONS   Let us remark that a marginal mutual information, e.g. $I(f_1 : f_2)$, does not serve as a perfect witness as it does not capture correlations in or with other observables. Only distributions from informationally complete measurements can be used in this sense. For marginal mutual informations there is an open conjecture resembling (11.14) (see [301])

$$I(f_1 : f_2) + I(g_1 : g_2) \overset{?}{\leq} I(\rho_1 : \rho_2). \tag{11.16}$$

Even is this relation is true, correlations between positions and momenta are not included. This shows once again the advantage of detecting a full phase space distribution.

11.2   GAUSSIAN STATES AND EXAMPLES

GAUSSIAN STATES   Let us start with a general (not necessarily pure) global Gaussian state ρ_{12} (see Section 3.4). Without loss of generality we assume that the expectation values of all quadratures vanish, which is justified by the translation invariance of all entropies. The global Husimi Q-distribution is of Gaussian from

$$Q_{12}(\vec{x}_1, \vec{k}_1, \vec{x}_2, \vec{k}_2) = \frac{1}{Z} e^{-\frac{1}{2}(\vec{x}_1, \vec{k}_1, \vec{x}_2, \vec{k}_2) C_{12} (\vec{x}_1, \vec{k}_1, \vec{x}_2, \vec{k}_2)^T}, \tag{11.17}$$

where we introduced a real, symmetric and positive definite matrix

$$C_{12} = \begin{pmatrix} C_1 & C_M \\ C_M^T & C_2 \end{pmatrix}, \tag{11.18}$$

which can be identified with the inverse covariance matrix of the global Husimi Q-distribution $C_{12} = V_{12}^{-1}$, and a normalization constant

$$Z = \sqrt{\det C_{12}}. \tag{11.19}$$

Recall that a pure Gaussian state is characterized by (3.67), i.e. has a covariance matrix $C_{12}^{-1} = V_{12} = \frac{1}{2}(\mathbb{1} + SS^T)$ for some symplectic matrix $S \in Sp(2N + 2M, \mathbb{R})$.

Reordering the exponent in (11.17) according to contributions from the two subsystems gives

$$\begin{aligned} &Q_{12}(\vec{x}_1, \vec{k}_1, \vec{x}_2, \vec{k}_2) \\ &= \frac{1}{Z} e^{-\frac{1}{2}(\vec{x}_1, \vec{k}_1) C_1 (\vec{x}_1, \vec{k}_1)^T - \frac{1}{2}(\vec{x}_2, \vec{k}_2) C_2 (\vec{x}_2, \vec{k}_2)^T - (\vec{x}_1, \vec{k}_1) C_M (\vec{x}_2, \vec{k}_2)^T}, \end{aligned} \tag{11.20}$$

showing that correlations (classical and quantum) between the two subsystems are encoded in the mixing term $(\vec{x}_1, \vec{k}_1) C_M (\vec{x}_2, \vec{k}_2)^T$. For $C_M = 0$, this term vanishes and the global Husimi Q-distribution factorizes into the two local Husimi Q-distributions.

Let us now compute the entropies. After a straightforward exercise in Gaussian integration one finds for the global Wehrl entropy

$$S(Q_{12}) = -\frac{1}{2} \ln \det C_{12} + N + M. \tag{11.21}$$

As the Wehrl entropy is maximized by a Gaussian Husimi Q-distribution for a given covariance matrix $V_{12} = C_{12}^{-1}$, the global WL inequality implies a second moment uncertainty relation in phase space

$$\det C_{12} \leq 1 \Leftrightarrow \det \left(\gamma_{12} + \frac{1}{2} \mathbb{1} \right) \geq 1. \tag{11.22}$$

Compared to the multi-dimensional Robertson-Schrödinger relation (cf. (3.75))

$$\det \gamma_{12} \geq \frac{1}{4^{N+M}}, \tag{11.23}$$

which is invariant under symplectic transformations and hence tight for all pure Gaussian states, our relation (11.22) is only tight for unsqueezed pure Gaussian states. This can be seen from the condition $S^T S = \mathbb{1}$, which needs to be fulfilled to reach equality in (11.22) and shows that S has to be an orthogonal matrix corresponding to rotations and displacements in phase space.

To compute the local entropies, we have to calculate the local Husimi Q-distributions first (we consider subsystem 2 in the following). Integrating out the complementary subsystem gives

$$Q_2(\vec{x}_2, \vec{k}_2) = \sqrt{\frac{\det C_{12}}{\det C_1}} \, e^{-\frac{1}{2}(\vec{x}_2, \vec{k}_2)\left(C_2 - C_M^T C_1^{-1} C_M\right)(\vec{x}_2, \vec{k}_2)^T}, \qquad (11.24)$$

leading to a local Wehrl entropy

$$S(Q_2) = -\frac{1}{2} \ln \det C_{12} + \frac{1}{2} \ln \det C_1 + M. \qquad (11.25)$$

Here, we employed the identity

$$\det C_{12} = \det C_1 \det \left(C_2 - C_M^T C_1^{-1} C_M\right), \qquad (11.26)$$

which holds true provided that $\det C_1 \neq 0$. Note that the local WL inequality implies

$$\det C_{12} \leq \det C_1. \qquad (11.27)$$

Now, we can compute the two quantities of our interest. We assume that all entropies are finite and obtain

$$S(Q_1 | Q_2) = N - \frac{1}{2} \ln \det C_1,$$
$$I(Q_1 : Q_2) = \frac{1}{2} \ln \frac{\det C_1 \det C_2}{\det C_{12}}. \qquad (11.28)$$

Surprisingly, the Wehrl conditional entropy is independent of the correlations encoded in C_M, which is in contrast to other classical (and also quantum) conditional entropies. Applying the conditional WL inequality leads to local uncertainty relations (11.10)

$$\det C_1 \leq 1, \quad \det C_2 \leq 1. \qquad (11.29)$$

Also, the necessary and sufficient separability criterion for pure states $I(Q_1 : Q_2) = 0$ translates into $C_M = 0$, as expected.

WEHRL MUTUAL INFORMATION AND SIMON CRITERION $C_M = 0$ is equivalent to $V_M = 0$ and $\gamma_M = 0$, which actually implies separability via the PPT criterion, see e.g. [44, 302, 303]. To show the converse, i.e. that $C_M = 0$ is also implied by the PPT criterion, we consider the Simon criterion (3.102). Hence, we restrict to $N = M = 1$ for simplicity in the following.

For a pure Gaussian state we can write $\gamma_{12} = \frac{1}{2}SS^T$ with $S \in Sp(4, \mathbb{R})$ and the multi-dimensional Robertson-Schrödinger relation (11.23) becomes tight $\det \gamma_{12} = \frac{1}{16}$. It is convenient to perform local symplectic transformations $S' = \mathrm{diag}(S_1, S_2)$ to write $\gamma_{12} = S' \Delta_{12} S'^T$ with Δ_{12} being the normal form of γ_{12} [44, 63]

$$\Delta_{12} = \begin{pmatrix} b & 0 & m_1 & 0 \\ 0 & b & 0 & m_2 \\ m_1 & 0 & c & 0 \\ 0 & m_2 & 0 & c \end{pmatrix}, \tag{11.30}$$

with $b, c, m_1, m_2 \in \mathbb{R}$. As symplectic transformations have unit determinant, we have $\det \gamma_{12} = \det \Delta_{12}$ and hence our first purity condition reads

$$\left(bc - m_1^2\right)\left(bc - m_2^2\right) = \frac{1}{16}. \tag{11.31}$$

We can find another purity condition for these four parameters by recalling that the symplectic eigenvalues λ_i (cf. (3.77)), which are all equal to $\frac{1}{2}$ for pure Gaussian states, do also appear (in $\pm$ pairs) in the matrix $i\Omega_{12}\Delta_{12}$. Here, $\Omega_{12} = \mathbb{1} \otimes \Omega$ where Ω denotes the symplectic metric for a single mode, cf. (3.58). One then finds the second purity condition

$$\det \gamma_1 + \det \gamma_2 + 2 \det \gamma_M = b^2 + c^2 + 2m_1 m_2 = \frac{1}{2}. \tag{11.32}$$

Note the appearance of the four invariants $I_1 = \det \gamma_1$, $I_2 = \det \gamma_2$, $I_3 = \det \gamma_M$ and $I_4 = \mathrm{Tr}(\Omega \gamma_1 \Omega \gamma_M \Omega \gamma_2 \Omega \gamma_M^T)$ under the local symplectic transformations S', which fulfill $\det \gamma_{12} = I_1 I_2 + I_3^2 - I_4$ [44].

Let us now apply the PPT criterion in phase space (3.100). It corresponds to a mirror reflection in one of the two local phase spaces and produces a new normal form covariance matrix Δ'_{12}. Assuming the PPT condition to be fulfilled, Δ'_{12} must be physical, i.e. fulfill both purity conditions. The local mirror reflection only affects $\det \gamma'_M = -\det \gamma_M$ and one can easily write down the second purity condition for Δ'_{12}

$$\det \gamma_1 + \det \gamma_2 - 2 \det \gamma_M = b^2 + c^2 - 2m_1 m_2 = \frac{1}{2}, \tag{11.33}$$

showing that the PPT criterion implies $\det \gamma_M = m_1 m_2 = 0$. Without loss of generality, we assume $m_2 = 0$ in the following. Then, combining (11.31) with (11.32) and using that $bc \geq 0$ leads to

$$f(b) = b^2 \left(\frac{1}{2} - b^2\right) \geq \frac{1}{16}. \tag{11.34}$$

However, the function $f(b)$ has a global maximum for $b = \frac{1}{2}$ with $f(b = \frac{1}{2}) = \frac{1}{16}$, and hence $b = c = \frac{1}{2}$ and $m_1 = m_2 = 0$, which finally implies $\gamma_M = C_M = 0$ for all pure separable Gaussian states. In this sense, the Wehrl mutual information being zero and the PPT criterion for Gaussian states, i.e. the Simon criterion (3.102), are equivalent.

TWO-MODE SQUEEZED VACUUM STATE We proceed with a more specific example, which is the Two-Mode Squeezed Vacuum (TMSV) state for $N = M = 1$ modes [63]

$$|\psi\rangle_{\text{TMSV}} = \sqrt{1 - \lambda^2} \sum_{n=0}^{\infty} (-\lambda)^n |n_1\rangle \otimes |n_2\rangle. \qquad (11.35)$$

Sometimes, the squeezing parameter $\lambda \in [0, 1]$ is also expressed in terms of another parameter $r \in [0, \infty)$, which are related via

$$\lambda = \tanh r. \qquad (11.36)$$

The TMSV state is rather special as it is a pure Gaussian state which is entangled for all $\lambda > 0$ while for $\lambda = 0$ it corresponds to two uncorrelated vacua. In the opposite limit $\lambda \to 1$ it becomes the maximally correlated EPR state first described in the famous paper [167]. Note that this state is the purification of the thermal state and hence both subsystems look thermal with a local temperature depending on the squeezing parameter λ.

The covariance matrix C_{12} is determined by

$$C_1 = C_2 = \mathbb{1}, \quad C_M = \text{diag}(\lambda, -\lambda), \qquad (11.37)$$

showing that the state is separable if and only if $\lambda = 0$ in which case the global Husimi Q-distribution factorizes into two vacua. Plugging the latter block matrices into (11.28) gives

$$S(Q_1|Q_2) = 1, \quad I(Q_1 : Q_2) = -\ln\left(1 - \lambda^2\right). \qquad (11.38)$$

We show the two quantities together with the quantum mutual information (which can be computed following [63, 304]) in Figure 11.1.
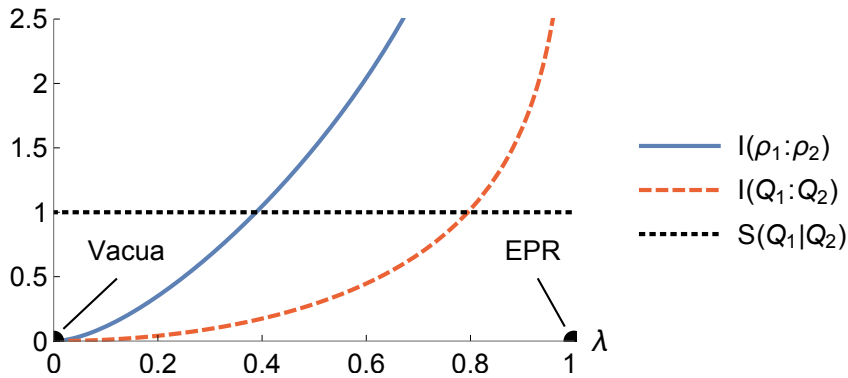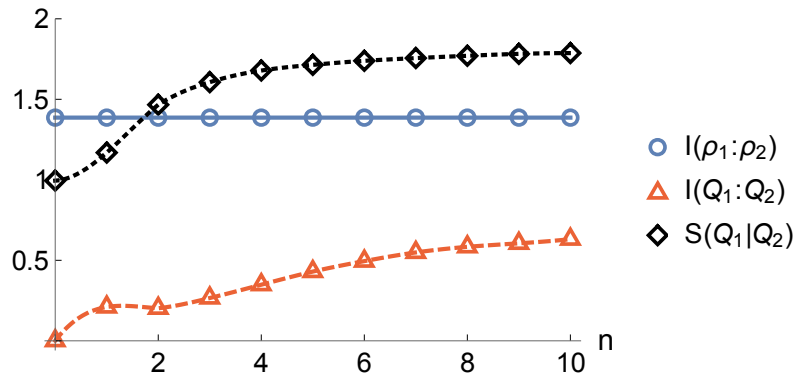

Figure 11.1: Wehrl entropic quantities $S(Q_1|Q_2)$ and $I(Q_1 : Q_2)$ together with the quantum mutual information $I(\rho_1 : \rho_2)$ for the TMSV state (11.35) as a function of the squeezing parameter $\lambda \in [0, 1]$.

Interestingly, the lower bound for the Wehrl conditional entropy is attained for all values of the squeezing parameter λ, despite the two subsystems exhibiting quantum correlations for $\lambda > 0$. This allows

for a straightforward application when trying to prepare the TMSV state (11.35). In a typical experiment, the precise value of λ fluctuates between subsequent realizations, which leads to a larger Wehrl conditional entropy $S(Q_1|Q_2) > 1$ as a consequence of conditional entropies being concave. Hence, any measured deviation from 1 shows additional (unwanted) classical correlations.

Furthermore, the Wehrl mutual information increases monotonically for increasing λ starting from zero for $\lambda = 0$ and becomes infinite for $\lambda \to 1$ showing that the EPR state exhibits a formally infinite amount of quantum correlations.

NOON STATES   At last, we consider a non-Gaussian example: the class of NOON states for $N = M = 1$ modes

$$|\psi\rangle_{\mathrm{NOON}} = \frac{1}{\sqrt{2(1 + \delta_{0n})}} \left(|n\rangle \otimes |0\rangle + |0\rangle \otimes |n\rangle \right), \qquad (11.39)$$

with $n \in \mathbb{N}$. Note that the normalization is chosen such that $|\psi\rangle_{\mathrm{NOON}} = |0\rangle \otimes |0\rangle$ for $n = 0$. A NOON state is separable if and only if $n = 0$. Note also that for $n = 1$ we obtain one of the four Bell states.

The global and local Husimi Q-distributions read

$$\begin{aligned}
Q_{12}(x_1, k_1, x_2, k_2) &= \frac{e^{-\frac{1}{2}\left(x_1^2 + k_1^2 + x_2^2 + k_2^2\right)}}{2^{n+1}n!(1 + \delta_{0n})} \\
&\times \left[(x_1 - ik_1)^n + (x_2 - ik_2)^n \right] \\
&\times \left[(x_1 + ik_1)^n + (x_2 + ik_2)^n \right], \\
Q_2(x_2, k_2) &= \frac{e^{-\frac{1}{2}\left(x_2^2 + k_2^2\right)}}{2^{n+1}n!} \left[\left(x_2^2 + k_2^2\right)^n + 2^n n! \right].
\end{aligned} \qquad (11.40)$$

The entropic Wehrl quantities have to be computed numerically, while the quantum mutual information evaluates to $I(\rho_1 : \rho_2) = 2\ln 2$ for all $n \in \mathbb{N}$. The results are shown in Figure 11.2.

Figure 11.2: Wehrl quantities $S(Q_1|Q_2)$ and $I(Q_1 : Q_2)$ together with the quantum mutual information $I(\rho_1 : \rho_2)$ for the NOON states (11.39) as a function of n. The curves represent interpolations.

It important to notice that no other entropic separability criterion is capable of witnessing the NOON states for all $n > 0$. The strong entropic criteria (which are also only valid for globally pure states) in [47] and [49] can only detect NOON states up to $n = 5$ and $n = 6$, respectively. In contrast, the Wehrl mutual information is positive for all $n > 0$ and hence it can witness all entangled NOON states. Also, the Wehrl mutual information increases monotonically for $n \geq 2$ and approaches the value of the quantum mutual information, which is independent of n.

12 | ENTROPIC SEPARABILITY CRITERIA

This chapter is taken from [C]. I proposed the project and S. F. and M. G. jointly supervised the work. I derived the entropic separability criteria and did all calculations. Also, I wrote the early drafts. The figures were produced by O. S., while I computed all quantities involved. All authors contributed in the writing of the manuscript.

For continuous variable systems, we have seen that the Wehrl mutual information provides a necessary and sufficient criterion for separability in the case of globally pure states. However, experiments have to deal with many imperfections in the form of additional classical correlations and also finite statistics. Although experimentally prepared states are often of high purity, they are never perfectly pure. To deal with this issue, we have to derive separability criteria, which are also valid for mixed separable states.

Instead of considering the global and local Husimi Q-distributions, we construct non-local variables similar to the setup discussed in Section 3.6. We adapt this strategy for the Husimi Q-distribution and derive new entropic separability criteria in Section 12.1. We apply these criteria to Gaussian states in Section 12.2 to obtain new second moment criteria and to study their behavior under symplectic transformations. Finally, we consider two classes of entangled non-Gaussian states, where other (entropic) criteria fail, in Section 12.3.

12.1 DERIVATION BASED ON ENTROPY POWER IN-EQUALITY

HUSIMI Q-DISTRIBUTION AND LOCAL ROTATIONS We start from a bipartition of $1 + 1$ modes each being described by continuous variables X_j and K_k fulfilling the Heisenberg algebra $[X_j, K_k] = i\delta_{jk}$ and with the subsystems being labeled by $j, k \in \{1, 2\}$. We allow every local variable pair (X_j, K_j) to be rotated in its local phase space by introducing angles $\vartheta_j \in [0, 2\pi)$, leading to rotated variables

$$\begin{pmatrix} R_j \\ S_j \end{pmatrix} = \begin{pmatrix} \cos\vartheta_j & \sin\vartheta_j \\ -\sin\vartheta_j & \cos\vartheta_j \end{pmatrix} \begin{pmatrix} X_j \\ K_j \end{pmatrix}. \tag{12.1}$$

The rotated variables still fulfill the Heisenberg algebra $[R_j, S_k] = i\delta_{jk}$. It is important to note that this seemingly technical step is necessary to end up with more general separability criteria. More precisely, the two angles ϑ_j provide optimization parameters. At this point, note also that

the angles may be selected *post-measurement*, which is an advantage of the Husimi Q-distribution over the marginal distributions $f(x)$ and $g(k)$, as the latter have to be measured along *preselected* axes. Furthermore, the local and global Wehrl entropies are invariant under such transformation as rotations have unit determinant, such that also the Wehrl mutual information and the Wehrl conditional entropy remain unmodified.

We associate local coherent states $|\alpha\rangle$ and $|\beta\rangle$ with the rotated variables via (cf. (3.45))

$$|\alpha\rangle = D(\alpha)\,|0\rangle \tag{12.2}$$

and similarly for $|\beta\rangle$. This allows to define the global Husimi Q-distribution as $Q_{12}(r_1, s_1, r_2, s_2) = (\langle\alpha| \otimes \langle\beta|)\rho_{12}(|\alpha\rangle \otimes |\beta\rangle)$. Note that one may equivalently start from $Q_{12}(x_1, k_1, x_2, k_2)$ and apply the two local rotations with ϑ_j.

NON–LOCAL VARIABLES   To derive separability criteria, we consider non-local EPR-type operators in the spirit of (3.104) (with $a = 1$)

$$R_\pm = R_1 \pm R_2, \quad S_\pm = S_1 \pm S_2, \tag{12.3}$$

which are designed to reveal information about the correlations between two local variables of the same type. Let us emphasize that their commutation relations are given by

$$[R_\pm, S_\pm] = 2i, \quad [R_\pm, S_\mp] = 0, \tag{12.4}$$

showing that the twisted assignment $(R_\pm, S_\mp)$ constitutes a set of commuting observables and that these variables are normalized differently as the local variables.

As the canonical phase space $\mathbb{R}^2$ is a vector space, the non-local operators fulfill eigenvalue equations just as their local counterparts, allowing us to work with the global phase space coordinates (r_+, s_+, r_-, s_-) instead of (r_1, s_1, r_2, s_2), which are related via

$$r_\pm = r_1 \pm r_2, \quad s_\pm = s_1 \pm s_2. \tag{12.5}$$

Transforming the global Husimi Q-distribution accordingly gives

$$\begin{aligned}
&Q_{12}(r_1, s_1, r_2, s_2) \\
\rightarrow &Q'_{12}(r_+, s_+, r_-, s_-) \\
=&\frac{1}{4}Q_{12}\left(\frac{r_+ + r_-}{2}, \frac{s_+ + s_-}{2}, \frac{r_+ - r_-}{2}, \frac{s_+ - s_-}{2}\right).
\end{aligned} \tag{12.6}$$

The Jacobi determinant of this transformation yields the prefactor $\frac{1}{4}$, which ensures that both global Husimi Q-distributions are normalized with respect to phase space measures of the same form

$$1 = \int \frac{\mathrm{d}r_1\,\mathrm{d}s_1}{2\pi} \frac{\mathrm{d}r_2\,\mathrm{d}s_2}{2\pi} Q_{12} = \int \frac{\mathrm{d}r_+\,\mathrm{d}s_+}{2\pi} \frac{\mathrm{d}r_-\,\mathrm{d}s_-}{2\pi} Q'_{12}. \tag{12.7}$$

Starting from the transformed distribution Q'_{12}, we integrate out two variables in a way that we end up with a phase space distribution of twisted variables

$$Q_\pm(r_\pm, s_\mp) = \int \frac{dr_\mp \, ds_\pm}{2\pi} \, Q'_{12}(r_+, s_+, r_-, s_-). \qquad (12.8)$$

Although we denote these two distributions with $Q_\pm$, they are *not* Husimi Q-distributions in a strict sense, as the underlying operators $R_\pm$ and $S_\mp$ commute instead of fulfilling the Heisenberg algebra. Nevertheless, $Q_\pm$ is normalized to one with respect to the standard phase space measure as a consequence of (12.7), non-negative and bounded from above by unity. Hence, we can still associate on entropy to it, in terms of which we will formulate our separability criteria.

SEPARABILITY CRITERIA FOR PURE STATES   We start with globally pure states, in which case separable states are given by product states of the form $\rho_{12} = \rho_1 \otimes \rho_2$ (cf. (3.89)). As usual, the global Husimi Q-distribution factorizes

$$Q_{12}(r_1, s_1, r_2, s_2) = Q_1(r_1, s_1) \times Q_2(r_2, s_2), \qquad (12.9)$$

which, after another variable transform, reveals that the marginalized distribution $Q_\pm$ reduces to a convolution

$$\begin{aligned}
Q_\pm(r_\pm, s_\mp) &= \int \frac{dr_1 \, ds_1}{2\pi} \, Q_1(r_1, s_1) \, Q_2(\mp r_1 \pm r_\pm, \pm s_1 \mp s_\mp) \\
&= \left(Q_1 * Q_2^{(\pm)} \right)(r_\pm, s_\mp),
\end{aligned} \qquad (12.10)$$

with the convention $Q_2^{(\pm)}(r, s) = Q_2(\pm r, \mp s)$.

This allows us to employ the entropic power inequality, which holds for all non-negative and normalized distributions irregardless of whether they are bounded from above or not. It provides a lower bound on the entropy of a convoluted distribution. Interestingly, the well-known fact that the variances of two Gaussian distributions are added up under a convolution is a special case of this inequality.

In two dimensions and applied to our setup the entropy power inequality reads [4, 56, 305]

$$e^{S(Q_\pm)} \geq e^{S(Q_1)} + e^{S(Q_2)}, \qquad (12.11)$$

where we used that entropies are invariant under mirror reflections, i.e. $S(Q_2^{(\pm)}) = S(Q_2)$. Thus, every pure product state needs to satisfy

$$S(Q_\pm) \geq \ln\left(e^{S(Q_1)} + e^{S(Q_2)} \right). \qquad (12.12)$$

Violation of the latter inequality for a globally pure state witnesses entanglement.

We can obtain a weaker and state-independent bound by applying the WL EUR to both local entropies, resulting in the two inequalities

$$S(Q_\pm) \geq 1 + \ln 2, \qquad (12.13)$$

which have to be fulfilled by all pure product states.

GENERALIZATION TO MIXED STATES The advantage of the weaker criteria (12.13) is that they can be generalized to mixed separable states of the form $\rho_{12} = \sum_i p(i)(\rho_1^i \otimes \rho_2^i)$ with $p(i) \geq 0$ and $\sum_i p(i) = 1$ (cf. (3.90)). For such states, the marginal distribution $Q_\pm$ obeys a similar decomposition

$$Q_\pm(r_\pm, s_\mp) = \sum_i p(i) Q_\pm^i(r_\pm, s_\mp). \tag{12.14}$$

Then, by using that any classical entropy is concave (cf. Section 2.2), we find that the weaker criteria (12.13) generalize identically to mixed states

$$S(Q_\pm) \geq \sum_i p(i) S(Q_\pm^i) \geq 1 + \ln 2, \tag{12.15}$$

The entropy of $Q_\pm$ is bounded from below by $1 + \ln 2$ for all separable states.

which are our general entropic criteria. Violation of these inequalities flags entanglement regardless of whether the state of interest is pure or mixed.

Note that one could also generalize the stronger criteria (12.12) in the same manner to

$$S(Q_\pm) \geq \sum_i p(i) \ln \left(e^{S(Q_1^i)} + e^{S(Q_2^i)} \right). \tag{12.16}$$

However, the resulting bound contains the mixing probabilities $p(i)$. Determining them is equally hard as determining whether the state is separable or not. Hence, these criteria are not useful in practice and we do not consider them any further.

12.2 SECOND MOMENT CRITERIA

DERIVING SECOND MOMENT CRITERIA Let us consider the class of Gaussian states. Without loss of generality, we assume all mean values to vanish. In this case, the twisted distribution $Q_\pm$ is of the form

$$Q_\pm(r_\pm, s_\mp) = \frac{1}{Z} e^{-\frac{1}{2}(r_\pm, s_\mp) V_\pm^{-1} (r_\pm, s_\mp)^T} \tag{12.17}$$

with

$$V_\pm = \gamma_\pm + \bar\gamma = \begin{pmatrix} \sigma_{r_\pm}^2 & \sigma_{r_\pm s_\mp} \\ \sigma_{r_\pm s_\mp} & \sigma_{s_\mp}^2 \end{pmatrix} + \mathbb{1} \tag{12.18}$$

denoting the corresponding covariance matrix and $Z = \det^{-1/2} V_\pm$ being a normalization constant. The diagonal entries of the covariance matrix $\gamma_\pm$ of the Wigner W-distribution are the standard variances of the marginal distributions $f(r_\pm)$ and $g(s_\mp)$, while the covariance $\sigma_{r_\pm s_\mp}$ is non-vanishing whenever the coordinate axes are not aligned with the principal axes. Note the absence of the prefactor $\frac{1}{2}$ in $\bar\gamma$ as

consequence of the different normalization of the non-local variables, see (12.4).

As the entropy $S(Q_\pm)$ is maximized by the Gaussian distribution (12.17) for a given covariance matrix $V_\pm$, we can infer a set of second moment criteria from (12.15), i.e.

$$1 + \frac{1}{2} \ln \det V_\pm \geq S(Q_\pm) \geq 1 + \ln 2, \tag{12.19}$$

implying

$$\det V_\pm \geq 4. \tag{12.20}$$

For Gaussian states, the criteria (12.20) and (12.15) are equivalent. In all other cases, the entropic criteria (12.13) are stronger than the second moment criteria (12.20) in the sense of (12.19).

BEHAVIOR UNDER SYMPLECTIC TRANSFORMATIONS   It is interesting to investigate how the separability criteria behave under symplectic transformations $S \in \mathrm{Sp}(2, \mathbb{R})$. While the Wigner W-distribution transforms as a scalar field under symplectic transformations (see e.g. [44]), the Husimi Q-distribution does not transform in a simple manner. This is due to the additional convolution with the Wigner W-distribution of the vacuum, which overweights specific widths and does not allow for a simple representation of squeezing transformations. Hence, we restrict the symplectic analysis to the second moment criteria, where the latter fact translates into the absence of invariance under squeezing.

We start from the covariance matrix of the Wigner W-distribution $\gamma_\pm$, which transforms as

$$\gamma_\pm \to \gamma'_\pm = S \gamma_\pm S^T, \tag{12.21}$$

such that its determinant is invariant $\det \gamma_\pm \to \det \gamma'_\pm = \det \gamma_\pm$ as $\det S = \det S^T = 1$.

In contrast, the covariance matrix of the Husimi Q-distribution transforms according to

$$\begin{aligned}
\det V_\pm \to \det V'_\pm &= \det \left(S \gamma_\pm S^T + \bar{\gamma} \right) \\
&= \det \left(\gamma_\pm + \bar{\gamma} \left(S^T S \right)^{-1} \right),
\end{aligned} \tag{12.22}$$

where we used that $\bar{\gamma} = \mathbb{1}$. This shows that the second moment criteria (12.20) are invariant under rotations and displacements, for which we have $S^T S = \mathbb{1}$, but not invariant under equal amounts of local squeezing described by $S = \mathrm{diag}(a, 1/a)$ with a being positive and real. Hence, the orientation of the axes in the twisted phase space is unimportant for entanglement witnessing when using a full phase space distribution.

COMPARISON TO MARGINAL SECOND MOMENT CRITERIA   In Section 3.6 we reviewed the two well-known MGVT (3.106) and DGCZ (3.107) separability criteria, which are also formulated for variances of non-local observables. As they are based on detecting marginal distributions and allow for squeezing in the local variables (cf. (3.104)), we reformulate our second moment criteria (12.20) for a comparison. Using (12.18) and (12.22) for $S = \mathrm{diag}(a, 1/a)$ we find

$$\left(\sigma_{r_\pm}^2 + a^2\right)\left(\sigma_{s_\mp}^2 + \frac{1}{a^2}\right) \geq 4 + \sigma_{r_\pm s_\mp}^2, \tag{12.23}$$

where we included the effect of squeezing with $a > 0$ being a non-negative real number.

To examine the tightness of the criteria (12.23) we minimize the left hand side over a. A unique minimum can be found for $a^2 = \frac{\sigma_{r_\pm}}{\sigma_{s_\mp}}$, in which case we end up with

$$\left(\sigma_{r_\pm}\sigma_{s_\mp} + 1\right)^2 \geq 4 + \sigma_{r_\pm s_\mp}^2. \tag{12.24}$$

When aligning the axes with the principles aces, for which $\sigma_{r_\pm s_\mp}^2 \geq 0$, we obtain

$$\sigma_{r_\pm}\sigma_{s_\mp} \geq 1, \tag{12.25}$$

which corresponds to the weak version of the MGVT criteria (3.106). This shows that after optimizing over the squeezing parameter a and the orientation of the distribution in phase space, our second moment criteria (12.23) and the weak MGVT criteria (3.106) are equivalent, which is depicted in Figure 12.1b.

This finding implies that the criteria (12.23) are *not* necessary for separability in the case of Gaussian states as they are strictly weaker then the full MGVT criteria (3.106) in general. Nevertheless, we can conclude that our entropic criteria (12.15) and the entropic WTSTD criteria (3.108) are equivalent in the Gaussian regime as both reduce to the weak MGVT criteria after optimization.

We compare the three weak second moment criteria (12.20), (3.106) and (3.107), i.e. for $a = 1$ and $\sigma_{r_\pm s_\mp} = 0$, in Figure 12.1a. In this case, our criteria are stronger than the DGCZ criteria, but weaker than the MGVT criteria.

Is is also interesting to note that the three criteria transform rather differently under symplectic transformations. Our criteria (12.23) and the weak MGVT criteria behave complementary with respect to rotations and squeezing: while our criteria are invariant under rotations, but not under squeezing (cf. (12.22)), it is the opposite for the weak MGVT criteria. In contrast, the weak DGCZ criteria are neither invariant under rotations, nor under squeezing.

(a) Comparison with weak MGVT and DGCZ criteria

(b) Equivalence to weak MGVT criteria after optimization
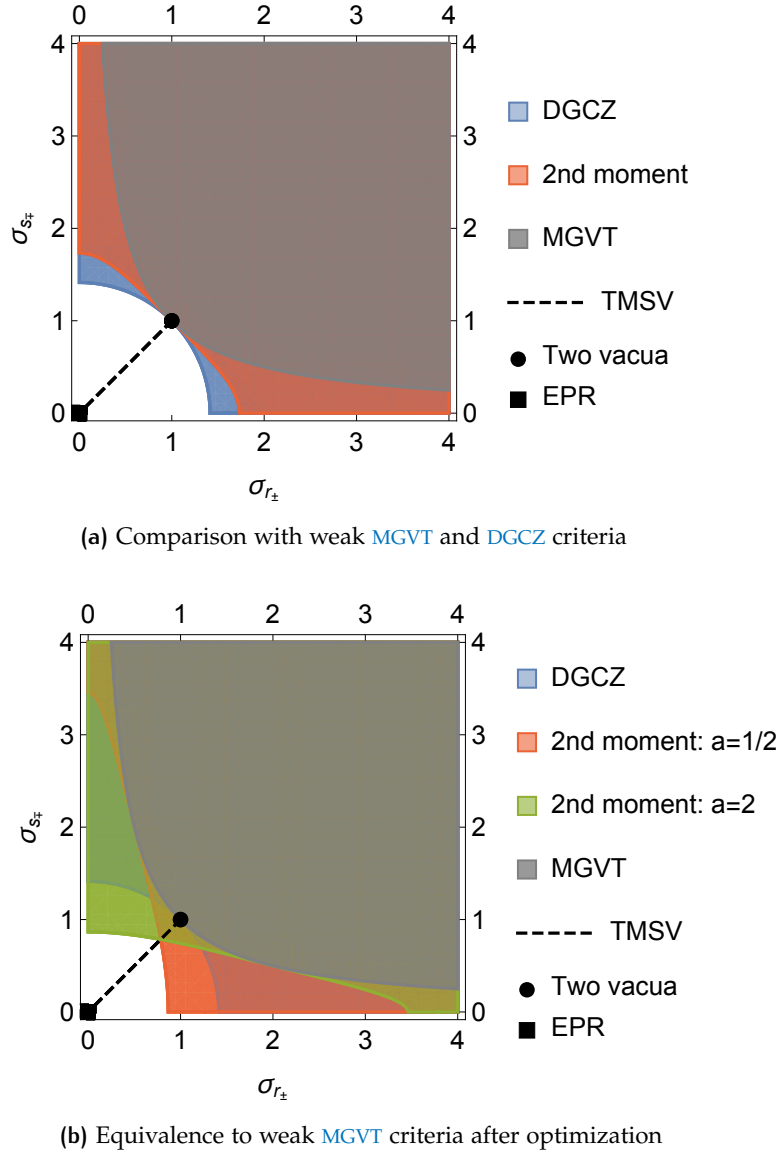
Figure 12.1: Weak DGCZ criteria (blue regions), second moment criteria
(12.23) (red and green regions) and MGVT criteria (gray regions)
as functions of the two marginal standard deviations $\sigma_{r_\pm}$ and
$\sigma_{s_\mp}$. The colored regions show where the criteria are fulfilled,
such that entanglement is certified below the curves. As a refer-
ence, we indicate the TMSV state (11.35) for all $\lambda \in [0,1]$ (black
dashed line) between the two vacua $\sigma_{r_\pm} = \sigma_{s_\mp} = 1$ (black dot)
and the maximally correlated EPR state $\sigma_{r_\pm} = \sigma_{s_\mp} = 0$ (black
square). In (a) we show all three criteria for $a = 1$ and $\sigma_{r_\pm s_\mp} = 0$.
The MGVT criteria are strongest, while the DGCZ are weakest.
Also, the curves touch for the two vacua. In (b) one can see that
optimizing our second moment criteria over a leads to the weak
MGVT criteria. More precisely, every point on the latter curve can
be touched by the curves corresponding to our second moment
criteria when adjusting a.

12.3 EXAMPLES

NOON STATES To benchmark the performance of our entropic criteria (12.15) in comparison with the entropic WTSTD criteria (3.108) in the non-Gaussian regime, we consider two non-trivial examples. We begin with the NOON states, which were investigated already with the Wehrl mutual information in Section 11.2 around (11.39). The corresponding distributions $Q_\pm(r_\pm, s_\mp)$ can be calculated analytically for $n \in \mathbb{N}$ in general, but we omit the corresponding expressions for brevity.

We recall that the NOON states can *not* be witnessed by any second moment criterion and that also the entropic WTSTD criteria based on marginal distributions fail. The strong version of the WTSTD entropic criteria for pure states (in analogy to our stronger criteria (12.12)) witnesses entanglement up to $n = 5$, while their Rényi-improved criterion works up to $n = 6$ [49].

Similarly, our entropic criteria (12.15) do also not certify any entanglement. However, the strong entropic criteria for pure states (12.12) witness entanglement up to $n = 11$ when using Q_+, which is shown in Figure 12.2 for $\vartheta_1 = \vartheta_2 = 0$.

We can conclude that existing entropic criteria are outperformed. Nevertheless, we recall that the Wehrl mutual information already provides a perfect witness for pure state entanglement and hence works even better than our stronger entropic criteria for pure states (12.12).
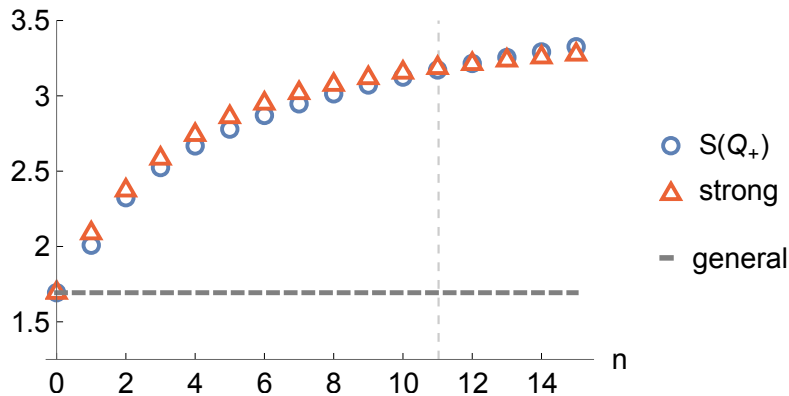
Figure 12.2: General (12.15) and strong entropic criteria for pure states (12.12) for the first fifteen NOON states. Entanglement is witnessed by the strong criteria up to $n = 11$, while the general criteria fail for all $n > 0$.

DEPHASED SCHRÖDINGER CAT STATE   As a true hardness test, we consider the dephased Schrödinger cat state, which is specified by the density operator [47]

$$
\begin{aligned}
\rho = N(r,s)\Big[& (|\alpha\rangle \otimes |\alpha\rangle)\,(\langle\alpha| \otimes \langle\alpha|) \\
& + (|-\alpha\rangle \otimes |-\alpha\rangle)\,(\langle-\alpha| \otimes \langle-\alpha|) \\
& - (1-z)\Big((|\alpha\rangle \otimes |\alpha\rangle)\,(\langle-\alpha| \otimes \langle-\alpha|) \\
& + (|-\alpha\rangle \otimes |-\alpha\rangle)\,(\langle\alpha| \otimes \langle\alpha|)\Big)\Big],
\end{aligned}
\tag{12.26}
$$

where $0 \le z \le 1$ is a parameter controlling its purity (the state is pure if and only if $z = 0$) and

$$
N(r,s) = \frac{1}{2}\left(1 + (1-z)\,e^{-2(r^2+s^2)}\right)
\tag{12.27}
$$

is a normalization constant. It is parameterized by the positions (r,s) (cf. also (3.46)) of the two local coherent states and for $\vartheta_1 = \vartheta_2 = 0$ its Husimi Q-distribution takes the form

$$
\begin{aligned}
Q(r_1,s_1,r_2,s_2) = N(r,s)\Big[& e^{-\frac{1}{2}\left((r-r_1)^2+(s-s_1)^2+(r-r_2)^2+(s-s_2)^2\right)} \\
& + e^{-\frac{1}{2}\left((r+r_1)^2+(s+s_1)^2+(r+r_2)^2+(s+s_2)^2\right)} \\
& + 2(1-z)\,e^{-r^2-s^2-\frac{1}{2}\left(r_1^2+s_1^2+r_2^2+s_2^2\right)} \\
& \times \cos\left(r(s_1+s_2) - s(r_1+r_2)\right)\Big].
\end{aligned}
\tag{12.28}
$$

As only the absolute value of α is of relevance in the twisted variables, we set $s = 0$ without loss of generality, in which case we can give an analytical expression (again for $\vartheta_1 = \vartheta_2 = 0$)

$$
\begin{aligned}
Q_\pm(r_\pm,s_\mp) = {} & \frac{e^{-\frac{1}{4}(r_\pm^2+s_\mp^2)}}{2\left(e^{2r^2}+1-z\right)} \\
& \times \begin{cases} 1 - z + e^{r^2}\cosh(r\,r_+) & \text{for} \quad (r_+,s_-), \\ e^{2r^2} + e^{r^2}(1-z)\cos(r\,s_+) & \text{for} \quad (r_-,s_+). \end{cases}
\end{aligned}
\tag{12.29}
$$

Despite this state being entangled for all $z < 1$ and $(r,s) \neq (0,0)$, it is notoriously hard to detect entanglement for the whole parameter range. Unsurprisingly, entanglement can not be witnessed for any choice of parameters by any second moment criterion. The entropic WTSTD criteria (3.108) witness entanglement at least for $r \gtrsim 2.5, s = 0$ and $z < 1$ when using $\vartheta_1 = \vartheta_2 = 0$ [47].

In contrast, our criteria witness entanglement for all values of $r > 0, s = 0$ and $z < 1$ when using Q_-, which is shown in Figure 12.3. Also, we see that our criteria (12.13) are violated most in the regime $0 \lesssim r \lesssim 1.5$, while for $r \gtrsim 2$ the violation becomes exponentially small.
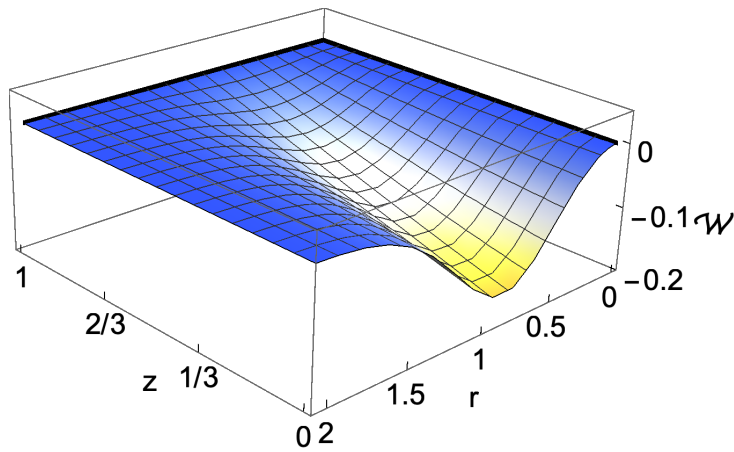
Figure 12.3: Entropic criteria (12.15) for the dephased Schrödinger cat state (12.26) as a function of the mixing parameter z and the position r. Entanglement is witnessed for all $z < 1$ and $r > 0, s = 0$, which is indicated by negative values for $\mathcal{W} = S(Q_-) - 1 - \ln 2$. The separable regions $z = 1$ and $r = 0$ are represented by black lines. The witness performs best in the region $0 \lesssim r \lesssim 1.5$ and converges to zero from below exponentially outside this region.

We can conclude that the dephased Schrödinger cat state is fully witnessed with our method, which is not possible with any other entropic separability criteria. Hence, the phase space approach turned out to be valuable for entanglement detection, which is why we amend our idea further in Chapter 13.

13 | GENERAL SEPARABILITY CRITERIA

The following discussion is based on [A, B]. The project covered in [A] was proposed by me and S. F. and M. G. jointly supervise it. I derived the general separability criteria and did all calculations. I wrote the current version of the manuscript. All authors participate in the ongoing discussion of the project.

We have seen that EURs lead to separability criteria which can be stronger than second moment criteria. In this chapter, we extend the ideas of Chapter 12, i.e. entropic separability criteria for a phase space distribution of twisted non-local variables, to arbitrary concave functionals of this distribution. The resulting criteria are rather general and have entropic and second moment criteria as special cases. As they come with the freedom to optimize over a large class of functions, the ability to detect entanglement is further improved, especially when applying them to discretized distributions.

We derive our general separability criteria in Section 13.1 and show that the previously derived criteria follow as special cases in Section 13.2. Finally, we discuss their experimental implementation and deduce their discretized analog in Section 13.3, which is applied to an example.

13.1 DERIVATION BASED ON PPT CRITERION

NON–LOCAL VARIABLES AND LIEB–SOLOVEJ THEOREM Following up on the discussion of non-local variables in Section 12.1, we start with the global transformed Husimi Q-distribution $Q'_{12}(r_+, s_+, r_-, s_-)$. Now, instead of integrating out two variables such that we obtain a distribution of twisted variables $Q_\pm(r_\pm, s_\mp)$, we consider the marginal distributions

$$Q'_\pm(r_\pm, s_\pm) = \int \frac{\mathrm{d}r_\mp \, \mathrm{d}s_\mp}{2\pi} Q'_{12}(r_+, s_+, r_-, s_-). \tag{13.1}$$

As the underlying measurement operators fulfill the Heisenberg algebra (cf. (12.4)), $Q'_\pm(r_\pm, s_\pm)$ are *true* Husimi Q-distributions.

As a consequence, $Q'_\pm(r_\pm, s_\pm)$ is constrained by the uncertainty principle. Most generally, the uncertainty principle is expressed by the Lieb-Solovej theorem [84]. Adapted to our situation, it states that for any concave function $f : [0, 1] \to \mathbb{R}$ with $f(0) = 0$ and any bipartite quantum state ρ_{12} the following inequality holds

$$\int \frac{\mathrm{d}r_\pm \, \mathrm{d}s_\pm}{2\pi} f\left(Q'_\pm\right) \geq \int \frac{\mathrm{d}r_\pm \, \mathrm{d}s_\pm}{2\pi} f\left(\bar{Q}'_\pm\right), \tag{13.2}$$

where

$$\bar{Q}'_{\pm}(r_{\pm}, s_{\pm}) = \frac{1}{2} e^{-\frac{1}{4}(r_{\pm}^2 + s_{\pm}^2)} \tag{13.3}$$

denotes the vacuum Husimi Q-distribution associated with the non-local operators $\boldsymbol{R}_{\pm}$ and $\boldsymbol{S}_{\pm}$. By translation invariance, one may equally well take any pure coherent state on the right hand side of (13.2).

Note that the right hand side of (13.2) evaluates to a constant as soon as the concave function f is fixed, while the left hand side is state-dependent. Note also that (13.2) reduces to the WL inequality, which in this case reads $S(Q'_{\pm}) \geq 1 + \ln 2$, for the choice $f(t) = -t \ln t$. In this sense, (13.2) is the generalization of an EUR in phase space to arbitrary concave functions.

PPT CRITERION IN PHASE SPACE In Section 3.6 we have already seen that uncertainty relations can be used to formulate separability criteria by employing the PPT criterion. With the rather general form of the uncertainty principle (13.2) at hand, let us apply the PPT criterion in phase space.

As seen in (3.100), the transposition in one of the two subsystems corresponds to a mirror reflection in the corresponding local phase space. To show this explicitly for the Husimi Q-distribution, let us consider two local bases $\{|n_1\rangle\}_n$ and $\{|n_2\rangle\}_n$ with $n \in \mathbb{N}$ denoting number eigenstates, such that any global state ρ_{12} may be written as

$$\rho_{12} = \sum_{n_1 n'_1 n_2 n'_2} p_{n_2 n'_2}^{n_1 n'_1} \left(|n_1\rangle \langle n'_1| \otimes |n_2\rangle \langle n'_2| \right) \tag{13.4}$$

with $p_{n_2 n'_2}^{n_1 n'_1} \in \mathbb{R}$ specifying the state. The corresponding global Husimi Q-distribution reads

$$Q_{12}(r_1, s_1, r_2, s_2) = \sum_{n_1 n'_1 n_2 n'_2} p_{n_2 n'_2}^{n_1 n'_1} \langle \alpha_1 | n_1 \rangle \langle n'_1 | \alpha_1 \rangle \langle \alpha_2 | n_2 \rangle \langle n'_2 | \alpha_2 \rangle \tag{13.5}$$

with the overlaps given by (cf. (10.6))

$$\langle \alpha_i | n_i \rangle = e^{-\frac{1}{2}(r_i^2 + s_i^2)} \frac{(r_i - i s_i)^{n_i}}{\sqrt{2^{n_i} n_i!}}. \tag{13.6}$$

Now, applying a partial transpose to subsystem 2 exchanges n_2 with n'_2, which transforms Q_{12} according to

$$Q_{12}(r_1, s_1, r_2, s_2) \xrightarrow{T_2} \sum_{n_1 n'_1 n_2 n'_2} p_{n_2 n'_2}^{n_1 n'_1} \langle \alpha_1 | n_1 \rangle \langle n'_1 | \alpha_1 \rangle \langle \alpha_2 | n'_2 \rangle \langle n_2 | \alpha_2 \rangle, \tag{13.7}$$

showing that the overlaps corresponding to subsystem 2 are replaced by their complex conjugates. Together with (13.6) this implies the mirror reflection property

$$Q_{12}(r_1, s_1, r_2, s_2) \xrightarrow{T_2} Q_{12}(r_1, s_1, r_2, -s_2), \tag{13.8}$$

as desired. Then, we can deduce that the marginals (13.1), which are true Husimi Q-distributions, transform as

$$Q'_\pm(r_\pm, s_\pm) \xrightarrow{T_2} Q_\pm(r_\pm, s_\mp),\tag{13.9}$$

showing the importance of the twisted variables in $Q_\pm(r_\pm, s_\mp)$ for entanglement witnessing.

GENERAL SEPARABILITY CRITERIA FOR CONCAVE FUNCTIONS The PPT criterion (3.98) tells us that any separable state has a positive partial transpose. Hence, $Q_\pm(r_\pm, s_\mp)$ has to be physical for every separable state and as such has to fulfill the uncertainty principle. In particular, it has to fulfill the Lieb-Solovej theorem (13.2).

Hence, every separable state satisfies

Any concave phase space functional of $Q_\pm$ is bounded from below for all separable states.

$$\int \frac{dr_\pm \, ds_\mp}{2\pi} f(Q_\pm) \geq \int \frac{dr_\pm \, ds_\mp}{2\pi} f(\bar{Q}_\pm)\tag{13.10}$$

for *any* concave $f : [0, 1] \to \mathbb{R}$ with $f(0) = 0$. These are our most general separability criteria in phase space. Violation of these inequalities for any f implies that ρ_{12} is entangled.

Note that for $f(t) = b\,t$ with $b \in \mathbb{R}$ being some parameter the relation (13.10) reduces to an equality for all states ρ_{12} as the distribution $Q_\pm$ is normalized. Note also that if the requirement $f(0) = 0$ is left out, both sides of the inequality (13.10) are either $+\infty$ or $-\infty$. However, also for $f(0) = 0$ it could still be that one side evaluates to $\pm\infty$.

13.2 SPECIAL FAMILIES OF CRITERIA

DERIVING FAMILIES OF ENTROPIC CRITERIA Let us show how separability criteria for entropic families arise. To that end, we apply a monotonically increasing function $g : \mathbb{R} \to \mathbb{R}$ to our criteria (13.10) to obtain

$$g\left(\int \frac{dr_\pm \, ds_\mp}{2\pi} f(Q_\pm)\right) \geq g\left(\int \frac{dr_\pm \, ds_\mp}{2\pi} f(\bar{Q}_\pm)\right)\tag{13.11}$$

for all separable states ρ_{12}. It is important to note that the function g does *not* strengthen or weaken the criteria (13.10) in the sense that neither more (or less) entangled states can be detected nor the signal-to-noise ratio can be improved (or worsened). Note also that one may allow g to be monotonically decreasing, in which case one has to choose f to be convex. In this case, the inequality (13.11) is *not* reversed.

We are particularly interested in quantities which are significant from an information theoretic perspective (see [12, 56] for overviews of families of classical entropies). For example, we obtain the family of Rényi-Wehrl entropies

$$S_\beta(Q_\pm) = \frac{1}{1 - \beta} \ln\left(\int \frac{dr_\pm \, ds_\mp}{2\pi} Q_\pm^\beta(r_\pm, s_\mp)\right)\tag{13.12}$$

when choosing $f(t) = t^\beta$ for $\beta \in (0,1) \cup (1,\infty)$ and $g(t) = \frac{\ln t}{1-\beta}$, which are concave/convex and monotonically increasing/decreasing functions for $\beta < 1/\beta > 1$, respectively. Then, the corresponding entropic separability criteria read

$$S_\beta(Q_\pm) \geq \frac{\ln \beta}{\beta - 1} + \ln 2. \tag{13.13}$$

In the limit $\beta \to 1$, the Rényi-Wehrl entropy (13.12) reduces to the ordinary Wehrl entropy, in which case (13.13) gives back the entropic criteria (12.13) discussed in Chapter 12. Alternatively, the same result can be obtained when choosing the concave function $f(t) = -t \ln t$ in (13.10).

With the same choice for $f(t) = t^\gamma$ and a modified $g(t) = \frac{1-t}{1-\gamma}$, we obtain another set of entropic criteria

$$S_\gamma(Q_\pm) \geq \frac{1}{1 - \gamma} \left(1 - \frac{2^{1-\gamma}}{\gamma} \right) \tag{13.14}$$

for Tsallis-Wehrl entropies

$$S_\gamma(Q_\pm) = \frac{1}{1 - \gamma} \left(1 - \int \frac{\mathrm{d}r_\pm \, \mathrm{d}s_\mp}{2\pi} \, Q_\pm^\gamma(r_\pm, s_\mp) \right) \tag{13.15}$$

with $\gamma \in (0,1) \cup (1,\infty)$. Note that the Tsallis-Wehrl entropies are also well-defined for $\gamma < 0$. However, in this case the condition $f(0) = 0$ is violated, such that we have to restrict to $\gamma > 0$ for the corresponding separability criteria.

DERIVING SECOND MOMENT CRITERIA One might ask whether the general criteria (13.10) imply stronger second moment criteria as (12.20), which were implied by our entropic criteria (12.13).

To that end, we consider Gaussian states, which are characterized by a distribution of the form (12.17). We start with the class of monomials $f(t) = \pm t^\beta$ for $\beta \lessgtr 1$, which (after applying suitable g functions) correspond to the Rényi-Wehrl and Tsallis-Wehrl entropic families. An analytical calculation leads to

$$\frac{1}{\beta} \left(\pm \det^{(1-\beta)/2} V_\pm \mp 2^{1-\beta} \right) \geq 0 \Leftrightarrow \det V_\pm \geq 4, \tag{13.16}$$

showing that we do not obtain an improvement over the already established second moment criteria (12.20).

We expect that this result generalizes to all concave functions f with $f(0) = 0$. While a proof is left for future work, this claim is supported by further numerical evidence. We have checked the functions $f(t) = \sin t, f(t) = \tanh t, f(t) = 1 - e^t$ and $f(t) = \ln(1 - t)$ and can report that in all cases we obtained the second moment criteria (12.20). One might speculate that this is related to the fact that the lower bounds in the Lieb-Solovej theorem (13.2) and the WL inequality (3.84) are attained by the same states.

13.3 EXPERIMENTAL APPLICATION

POSSIBLE EXPERIMENTAL REALIZATIONS So far we have mainly focused on developing separability criteria and the discussion has been a rather theoretical one. Let us now comment on possible ways to detect the Husimi Q-distribution and comment on how finite measurement accuracy affect the various separability criteria we have derived throughout Part iv.

The measurement of the Husimi Q-distribution has been accomplished for several experimental setups. In principle, one can either employ a tomographic approach by displacing the state of interest and projecting it onto the vacuum afterwards or one can perform a heterodyne measurement. These strategies have been successfully applied in quantum optics [306–309]. Rather recently, similar techniques were used for Bose-Einstein condensates [310, 311], cold atoms in optical cavities [312, 313] and circuit quantum electrodynamics architectures [314]. Besides the method used in [311], the listed works were concerned with monopartite Husimi Q-distributions. Hence, extensions to bipartite systems are required, which is an ongoing challenge.

In any case, one of the main advantages of detecting a full phase space distribution over two marginal distributions is that one does not have to preselect the local angles ϑ_i, such that angle tomography, which requires a lot of experimental runs, is avoided. Also, especially in cold atoms systems, these angles are difficult to control. On top of that, our phase space criteria showed great performance.

DISCRETIZATION IN PHASE SPACE For an experimental application of the general separability criteria (13.10) (or any other of our criteria) it is important to consider experimental limitations. Although we deal with continuous phase space variables (r, s), any measurement procedure is fundamentally limited, for example by the precision of a detector or finite statistics effects. In particular, an experiment will produce a discrete empirical sample distribution or (possibly after binning) a histogram for discrete probabilities associated with bins instead of a smooth probability density function.

Discretizing distributions corresponds to coarse-graining and hence it is not surprising that the entropy (or any other concave phase space functional) grows. Therefore, discretized witnesses are generically weaker than their continuous analogs, excluding the possibility of false-positive results. However, finite statistics can cause the entropy of a (possibly discretized) distribution to be under- or overestimated, which is why in this case the proper inclusion of statistical errors becomes important. These effects are beyond our scope for the following discussion, but have to be taken into account when applying *any* entanglement witness in an actual experiment.

In the following, we adapt our criteria (13.10) for discrete histograms (see also [50, 113]). We start with discretizing phase space into bins with discrete coordinates (at the center of each bin)

$$r_i = i \, \Delta r_i, \quad s_j = j \, \Delta s_j, \tag{13.17}$$

with $i, j \in \mathbb{Z}$ labeling these discrete coordinates and $\Delta r_i \Delta s_j$ being the phase space area element associated with the (r_i, s_j)-th bin. Note that with this construction we allow for adaptive bin sizes.

To every bin we associate a discrete probability by integrating the Husimi Q-distribution over the corresponding bin

$$q(r_i, s_j) = \frac{1}{2\pi} \int_{\Delta r_i(i-1/2)}^{\Delta r_i(i+1/2)} \mathrm{d}r \int_{\Delta s_j(i-1/2)}^{\Delta s_j(i+1/2)} \mathrm{d}s \, Q(r,s). \tag{13.18}$$

This distribution is normalized to unity as an ordinary probability distribution

$$1 = \sum_{i,j=-\infty}^{\infty} q(r_i, s_j) \tag{13.19}$$

and resembles the continuous Husimi Q-distribution in the sense

$$\frac{2\pi}{\Delta r_i \, \Delta s_j} q(r_i, s_j) \to Q(r,s), \tag{13.20}$$

in the continuum limit $\Delta r_i \, \Delta s_j \to 0$.

Additionally, we define another distribution

$$Q(r,s;r_i,s_j) = \begin{cases} \frac{Q(r,s)}{q(r_i,s_j)} & \text{for} \quad (r,s) \in \left(\Delta r_i(i \pm 1/2), \Delta s_j(j \pm 1/2) \right), \\ 0 & \text{else,} \end{cases} \tag{13.21}$$

which is the distribution over (r,s) conditioned on measuring inside the (r_i, s_j)-th bin. In simple words, it corresponds to the averaged Husimi Q-distribution over this bin. It is normalized to unity with respect to the standard phase space measure

$$1 = \frac{1}{2\pi} \int_{\Delta r_i(i-1/2)}^{\Delta r_i(i+1/2)} \mathrm{d}r \int_{\Delta s_i(i-1/2)}^{\Delta s_i(i+1/2)} \mathrm{d}s \, Q(r,s;r_i,s_j). \tag{13.22}$$

We are interested in discretizing expressions of the form $\int \frac{\mathrm{d}r \, \mathrm{d}s}{2\pi} f(Q)$ for functions f fulfilling our usual conditions. To that end, we plug in (13.21) and break up the integral into integrals over single bins

$$\int \frac{\mathrm{d}r \, \mathrm{d}s}{2\pi} f(Q) = \sum_{i,j=-\infty}^{\infty} \frac{1}{2\pi} \int_{\Delta r_i(i-1/2)}^{\Delta r_i(i+1/2)} \mathrm{d}r \int_{\Delta s_j(i-1/2)}^{\Delta s_j(i+1/2)} \mathrm{d}s$$
$$\times f\left(q(r_i, s_j) \, Q(r,s;r_i,s_j) \right). \tag{13.23}$$

For every single integral we can use that the uniform distribution $Q(r, s; r_i, s_j) = \frac{2\pi}{\Delta r_i \Delta s_j}$ is majorized by all others, i.e., in simple words, is most spread out over a single bin [121]. As f is concave, we obtain an upper bound for our quantity of interest in terms of the discretized Husimi Q-distribution (13.18)

$$\int \frac{dr\, ds}{2\pi} f(Q) \le \sum_{i,j=-\infty}^{\infty} \frac{\Delta r_i \Delta s_j}{2\pi} f\left(\frac{2\pi}{\Delta r_i \Delta s_j} q\right). \tag{13.24}$$

EXAMPLE: WEHRL ENTROPY Exemplary, we consider the Wehrl entropy defined via $f(t) = -t \ln t$ for equal bins $\Delta r_i \Delta s_j \equiv \Delta r \Delta s$, such that (13.24) evaluates to

$$S(Q) \le S(q) + \ln \frac{\Delta r \Delta s}{2\pi}. \tag{13.25}$$

Equality is reached in the continuum limit $\Delta r \Delta s \to 0$, in accordance with the definition of the differential entropy (2.11) in Section 2.2.

(a) $\bar{q}(r_i, s_j)$ for $\Delta r \Delta s = 0.5$

(b) $\bar{q}(r_i, s_j)$ for $\Delta r \Delta s = 1.5$

(c) Wehrl entropy bound (13.25) for varying $\Delta r \Delta s$
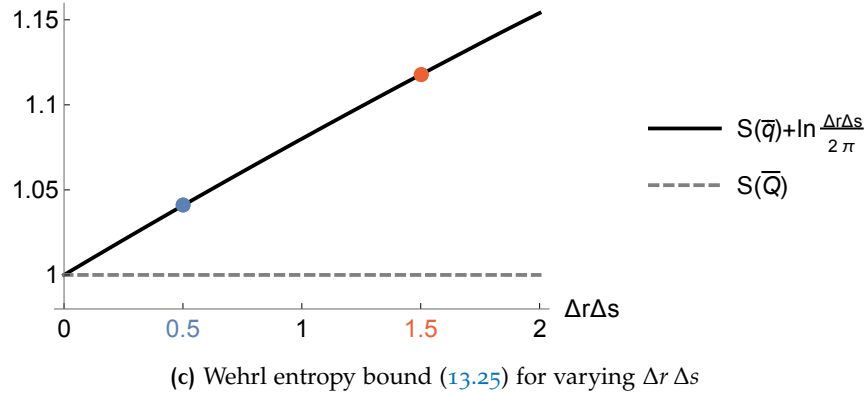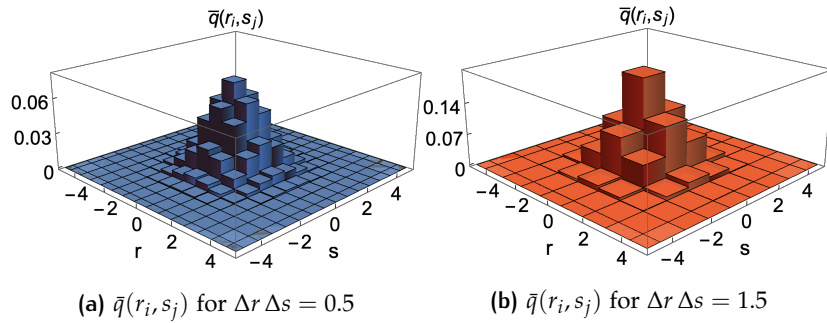
Figure 13.1: We depict the discretized vacuum Husimi Q-distribution $\bar{q}(r_i, s_j)$ for $\Delta r \Delta s = 0.5$ in (a) and for $\Delta r \Delta s = 1.5$ in (b). Both sides of the entropic inequality (13.25) are shown as a function of the phase space area element $\Delta r \Delta s$ in (c). The right hand side is approximately linear in this regime, which ceases to be the case for sufficiently large $\Delta r \Delta s$ (not shown in the figure). The blue and red points correspond to the distributions in (a) and (b), respectively. The relation (13.25) becomes tight in the continuum limit $\Delta r \Delta s \to 0$.

We show both sides of the latter inequality as a function of the discrete area element $\Delta r \, \Delta s$ for the vacuum in Figure 13.1c together with two exemplary discretized Husimi Q-distributions $\bar{q}(r_i, s_j)$ in Figure 13.1a (with $\Delta r \, \Delta s = 0.5$) and Figure 13.1b (with $\Delta r \, \Delta s = 1.5$).

DISCRETIZED UPPER BOUNDS ON SEPARABILITY CRITERIA   The previous considerations allow us to formulate (weaker) separability criteria for a discrete distribution $q_\pm$ obtained from $Q_\pm$ in the sense of (13.18). Applying (13.24) to our general separability criteria (13.10) results in the discretized criteria

$$\sum_{i,j=-\infty}^{\infty} \frac{\Delta r_{\pm,i} \, \Delta s_{\mp,j}}{2\pi} f\left(\frac{2\pi}{\Delta r_{\pm,i} \, \Delta s_{\mp,j}} q_\pm \right) \geq \int \frac{\mathrm{d}r_\pm \, \mathrm{d}s_\mp}{2\pi} f(\bar{Q}_\pm), \quad (13.26)$$

which are fulfilled for all separable states (3.90) and all concave f with $f(0) = 0$. Note that if one works with sampled data, in which case the bin sizes are not fixed a priori, one can optimize the left hand side over the bin sizes $\Delta r_{\pm,i} \, \Delta s_{\mp,j}$.

To end up with discretized versions of entropy functionals on the left hand side of (13.26), it is necessary to assume equal bin sizes $\Delta r_{\pm,i} \, \Delta s_{\mp,j} \equiv \Delta r_\pm \, \Delta s_\mp$. For example, we find the chain of inequalities

$$S_\beta(q_\pm) + \ln \frac{\Delta r_\pm \, \Delta s_\mp}{2\pi} \geq S_\beta(Q_\pm) \geq \frac{\ln \beta}{\beta - 1} + \ln 2 \qquad (13.27)$$

for the family of discretized Rényi-Wehrl entropies

$$S_\beta(q_\pm) = \frac{1}{1-\beta} \ln \left(\sum_{i,j=-\infty}^{\infty} q_\pm^\beta(r_{\pm,i}, s_{\mp,j}) \right). \qquad (13.28)$$

In contrast, one cannot write down criteria for discretized Tsallis-Wehrl entropies in simple form, as the latter result was established using that the logarithm of a product equals the sum of two logarithms.

DEPHASED SCHRÖDINGER CAT STATE   We examine the tightness of the discretized separability criteria for the family of Rényi-Wehrl entropies (13.27) and the dephased Schrödinger cat state (12.28). To that end, we consider the values $s = z = 0$ and vary $r \in [0, 2]$, which corresponds to the front slice in Figure 12.3.

In this regime, our entropic criteria (12.13) witnessed entanglement, while all other entropic and second moment criteria failed. We show the continuous and discretized Rényi-Wehrl criteria (13.13) and (13.27), respectively, for $\beta = 1$ (corresponding to the ordinary Wehrl entropy) in Figure 13.2a, for $\beta = 2$ in Figure 13.2b and for $\beta = 10$ in Figure 13.2c.

Detection of entanglement is (depending on the bin size $\Delta r_- \, \Delta s_+$) prevented for sufficiently small and large r, which is inherent with discretization. However, the possibility of optimizing over β can reduce this effect noticeably indicating the potential of the general discrete

separability criteria (13.26). In particular, one might prefer to choose large values for β, such that small values of $q_{\pm}(r_{\pm,i}, s_{\mp,j})$, which are hard to detect, are suppressed.

(a) $\beta = 1$
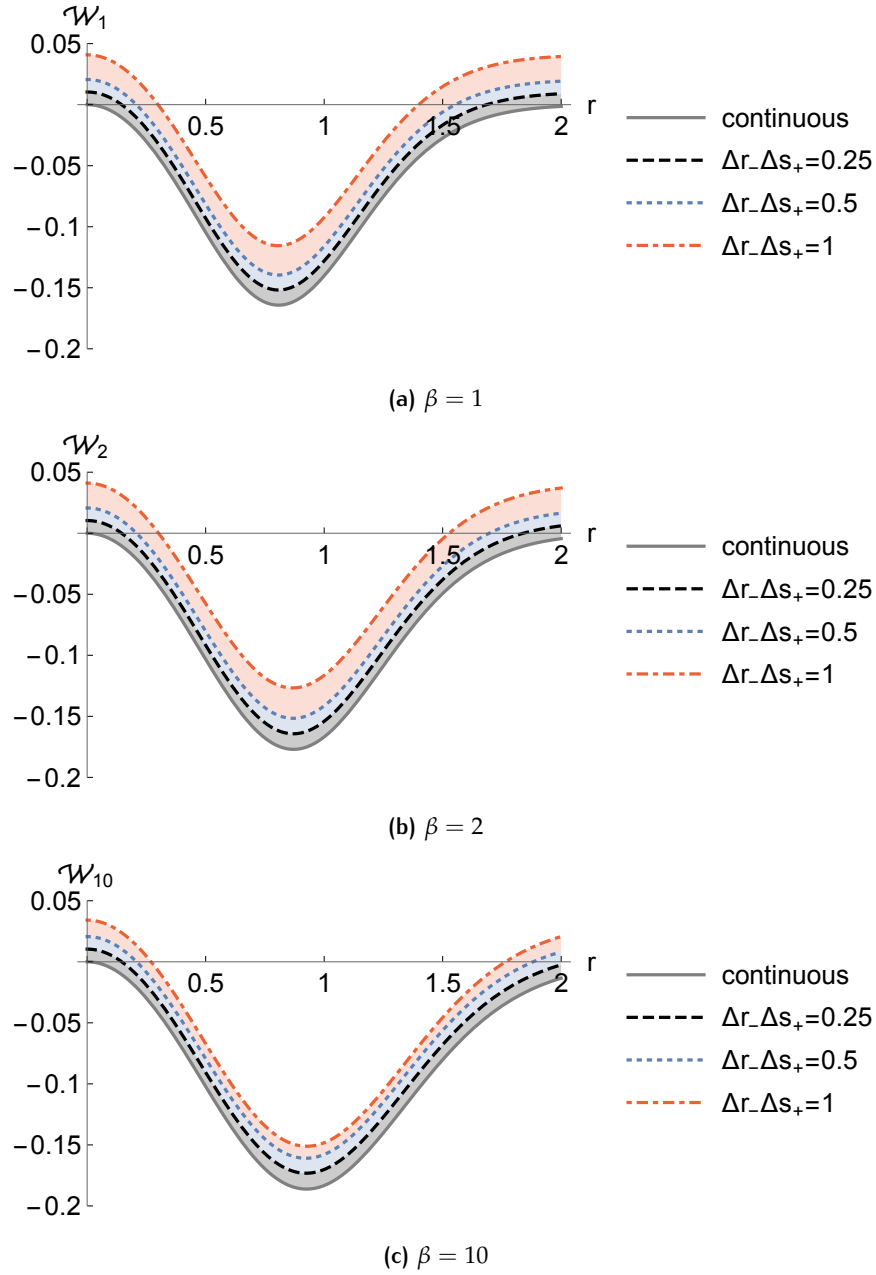

(b) $\beta = 2$

(c) $\beta = 10$

Figure 13.2: Continuous vs. discretized Rényi-Wehrl entropic criteria (cf. (13.27)) for the dephased Schrödinger cat state (12.26) with $s = z = 0$ as a function of r for various bin sizes $\Delta r_- \Delta s_+$. Entanglement is witnessed for negative $\mathcal{W}_\beta$, which is defined as the difference of the two sides in (13.27). We show the Wehrl criteria in (a), while (b) and (c) depict higher-order Rényi-Wehrl entropies with $\beta = 2$ and $\beta = 10$, respectively. For discretized data, detecting entanglement becomes harder for small $r \lesssim 0.25$ and large $r \gtrsim 1.5$.

14 | CONCLUSION AND OUTLOOK

GENERAL PERSPECTIVE   We have used relative entropy and entropy in the quantum mechanical phase space to describe equilibrium, uncertainty and entanglement. In doing so, especially the concept of relative entropy turned out to be an all-purpose tool. From our various analyses, it is tempting to believe that *distinguishability* may be even more fundamental than *information*. In any case, many more investigations are needed to formulate physical laws with the methods of information theory. For this purpose, relative entropy will certainly be one of the fundamental tools. In the following, we discuss our contributions successively and formulate open problems.

EQUILIBRIUM   In Chapter 5, we formulated a principle of inference in terms of relative entropy for classical distributions and quantum states. We have shown that it correctly predicts the uniform distribution and the maximally mixed state as priors if no further information is given. We demonstrated that relative entropy can be replaced by a true distance measure, the Hellinger distance, allowing us to conclude that the optimal prior corresponds to the central point on the manifold of allowed distributions for a simple example. However, there are three main open problems.

First, the complexity of the integral over allowed distributions or states prevented us from implementing further constraints in a straightforward way. Although we have shown that the optimal prior is always given by the expectation value over all allowed distributions or states, it is left to provide methods to evaluate our principle in general cases.

Second, for continuous variables the aforementioned integral becomes a functional integral, complicating the integration procedure even more. In this case, one should at least try to confirm that the uniform prior is correctly predicted when considering a finite interval.

Third, it is interesting to check how generic the functional used for comparing two distributions can be. It could be that all statistical divergences are well-suited or even that only convexity with respect to the argument which is integrated out is sufficient to arrive at a satisfactory principle of inference.

In Chapter 6, we developed thermodynamics in terms of thermodynamic reference states and relative entropy. This description goes beyond equilibrium situations as the thermodynamic variables temperature and chemical potential can be associated also with nonequilibrium states. In this formulation the second law of thermodynamics is a simple consequence of the monotonicity property of quantum

relative entropy when assuming that thermodynamic reference states are left invariant by time evolution.

Here, it is of special interest to see whether the principle of minimum expected relative entropy is capable of predicting the canonical state for a given energy expectation values along the lines of Section 3.3. Also, a more detailed investigation regarding the CPTP and PTP maps which leave the thermodynamic reference state invariant is in order. In particular, finding an explicit form of the Kraus operators and checking which other states are left invariant is desirable.

In Chapter 7, we derived second law-like inequalities for relativistic fluids in the context of local QFT. We made use of relative (entanglement) entropy rendering all involved quantities finite and well-defined.

As our discussion was rather conceptual, a more concrete formalism needs to be developed. In particular, one has to find general ways to compute relative entanglement entropies in QFTs, which is an ongoing effort [272]. Also, our construction made use of the double light cone, which is of course a simplifying assumption. It is of great importance to generalize our ideas to general open systems. What we have in mind here is a local version of the Schwinger-Keldysh formalism in QFT, which is also a current project [273]. Such a formalism would allow us to study the time evolution of a local region influenced by the interaction with its complement. It would be of great interest to verify our second law-like relations in such a scenario, at least for some concrete examples.

UNCERTAINTY   In Chapter 8, we formulated the uncertainty principle in terms of relative entropy. Our REUR holds naturally for discrete as well as continuous variables. In particular, it keeps its form in the continuum limit.

For future work it is of particular interest to generalize this ansatz to bipartite systems where one system acts as (quantum) memory. The corresponding EURs are typically formulated in terms of conditional entropy and can be applied for example for entanglement witnessing. A formulation in terms of conditional relative entropy is desirable as it would unify existing discrete and continuous variable relations.

In Chapter 9, we extended the concept of entropic uncertainty to QFT culminating in the field-theoretic REUR. We argued that relative entropy is the natural measure for entropic uncertainty in field theories and showed that the relation remains divergence-free in the continuum limit. In particular, we considered a free particle and concluded that its entropic uncertainty agrees with the one of a single mode.

As our work marks the first step in this direction, it brings a lot of questions and new problems.

First, it is interesting to investigate interacting theories, for example a scalar field with a ϕ^4 interaction term. We expect that our argu-

ment still works and that the resulting REUR acquires higher-order correlation functions in the bound.

Second, one might study other field theories and corresponding EURs as starting points. A formulation of the WL inequality for a scalar quantum field has been worked out in [315]. One might also try to find a formulation for spin degrees of freedom starting from the MU relation, which is of interest for experimental purposes. Another interesting type of theories are gauge theories as in this case the functional integral is overcounting the physical field configurations, such that a gauge fixing procedure is required.

Third, it is highly interesting to investigate how a field-theoretic REUR can be used to constrain entanglement in field theories. It might be easiest to analyze the Wehrl and quantum mutual informations in this regard as every mutual information is UV-finite in the continuum limit, but also formulations in terms of conditional relative entropy might be possible.

In Chapter 10, we investigated EURs in phase space for various classes of states. From our analytical and numerical evidence we conjectured that the WL relation is closest to equality almost everywhere.

Of course, it is natural to attempt to prove this conjecture. Let us also recall that so far there is no proven EUR in phase space which is tight for all pure Gaussian states. If such a relation is found, it might itself be closer to equality than the WL relation almost everywhere.

Furthermore, one may conduct a similar comparison for quantum spin systems described by an $SU(2)$ algebra. In particular, one may analyze the tightness of the MU relation and an EUR in phase space. For quantum spins, the corresponding phase space has the geometry of a 2-sphere embedded into the space spanned by the three spatial components of the angular momentum operator $\boldsymbol{J}$. Using spin coherent states, a Husimi Q-distribution can be defined on this sphere, whose entropy is constrained by an analog of the WL inequality.

ENTANGLEMENT   In Chapter 11, we investigated the Wehrl conditional entropy and the Wehrl mutual information for the canonical phase space. For the former we found an EUR strengthening the well-known monotonicity property of the Wehrl entropy, while the latter turned out to be a perfect witness for entanglement and a lower bound on the entanglement entropy for pure states. As such, it was able to witness the whole class of NOON states, which is not possible with any other entropic criterion.

As pointed out already, it is of interest to extend these concepts to QFT. More importantly, one should try to generalize these ideas to other observables characterized by other algebras. Here, the crucial advantage of the phase space approach may become visible. For example, the mutual information of a full phase space distribution will

always be a perfect witness for pure state entanglement, see e.g. [316] for a study of spin observables described by a $SU(2)$ algebra.

In Chapter 12, we constructed entropic separability criteria based on the Husimi Q-distribution. These criteria circumvented angle tomography and outperformed known entropic criteria, which we exemplified with the dephased Schrödinger cat state. Our entropic criteria were the first being capable of witnessing entanglement for this state for the full parameter ranges.

Also here generalizations to other observables are most interesting. Note that existing entropic separability criteria for discrete and continuous variable systems can not be related due to the divergences the classical entropy exhibits in the continuum limit. In contrast, a phase space distribution is always continuous and it is rather the geometry of the phase space, which depends on the observables under consideration.

Hence, the phase space approach would not only lead to new entropic criteria for other systems, but would also constitute a system independent approach to entropic separability criteria. Note that for such generalizations it is probably more convenient to start from the PPT criterion and apply a suitable EUR in phase space, similarly to the argument presented in Chapter 13, instead of trying to generalize the entropy power inequality.

In Chapter 13, we derived general separability criteria for concave functions and generalized them to discretized distributions. Especially the latter criteria should be of high interest to the experimental community. These criteria leave us with the freedom to optimize over the set of concave functions, allowing us to witness entanglement where our discretized entropic crtieria failed.

For future work, it is most interesting to apply these criteria in an actual experiment. In particular, one might try to certify entanglement where other criteria fail, for example because of too poor resolution, which may be circumvented by choosing suitable concave functions.

Also, a generalization to quantum spins with $SU(2)$ algebra should be striven for as an analog of the Lieb-Solovej theorem has been proven already in this case [85].

BIBLIOGRAPHY

[1] W. Heisenberg, *Physics and Philosophy: The Revolution in Modern Science* (New York: Harper, 1958).

[2] J. A. Wheeler, "Information, Physics, Quantum: The Search for Links," in *Proceedings III International Symposium on Foundations of Quantum Mechanics* (1989), pp. 354–358.

[3] C. E. Shannon, "A mathematical theory of communication," Bell Syst. Tech. J. **27**, 379–423 (1948).

[4] C. E. Shannon, "A mathematical theory of communication," Bell Syst. Tech. J. **27**, 623–656 (1948).

[5] L. Boltzmann, *Vorlesungen über Gastheorie*, 1896.

[6] J. W. Gibbs, "On the Equilibrium of Heterogeneous Substances," Trans. Conn. Acad. Arts Sci. **3**, 300–320 (1879).

[7] J. W. Gibbs, *Elementary principles in statistical mechnics* (Yale University Press, 1902).

[8] J. von Neumann, *Mathematical Foundations of Quantum Mechanics* (Princeton University Press, 1955).

[9] L. Szilard, "Über die Entropieverminderung in einem thermodynamischen System bei Eingriffen intelligenter Wesen," Z. Phys. **53**, 840–856 (1929).

[10] R. Landauer, "Information is Physical," Phys. Today **44**, 23–29 (1991).

[11] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary Edition* (Cambridge University Press, 2010).

[12] M. M. Wilde, *Quantum Information Theory* (Cambridge University Press, 2013).

[13] I. Bengtsson and K. Zyczkowski, *Geometry of Quantum States, 2nd ed.* (John Wiley and Sons, 2017).

[14] R. Jozsa and N. Linden, "On the Role of Entanglement in Quantum-Computational Speed-Up," Proc. Math. Phys. Eng. Sci. **459**, 2011–2032 (2003).

[15] T. D. Ladd, F. Jelezko, R. Laflamme, Y. Nakamura, C. Monroe, and J. L. O'Brien, "Quantum computers," Nature **464**, 45–53 (2010).

[16] D. Marolf, "The black hole information problem: past, present, and future," Rep. Prog. Phys. **80**, 092001 (2017).

[17] J. Polchinski, "The Black Hole Information Problem," in *New Frontiers in Fields and Strings* (2016) Chap. 6, pp. 353–397.

[18] D. Harlow, "Jerusalem lectures on black holes and quantum information," Rev. Mod. Phys. **88**, 015002 (2016).

[19] S. Kullback and R. A. Leibler, "On Information and Sufficiency," Ann. Math. Stat. **22**, 79–86 (1951).

[20] S. Kullback, *Information Theory and Statistics* (Dover Publications, 1968).

[21] H. Umegaki, "Conditional expectation in an operator algebra. IV. Entropy and information," Kodai Math. Semin. Rep. **14**, 59–85 (1962).

[22] A. Wehrl, "General properties of entropy," Rev. Mod. Phys. **50**, 221–260 (1978).

[23] A. Wehrl, "On the relation between classical and quantum-mechanical entropy," Rep. Math. Phys. **16**, 353–358 (1979).

[24] E. T. Jaynes, "Information Theory and Statistical Mechanics," Phys. Rev. **106**, 620–630 (1957).

[25] E. T. Jaynes, "Information Theory and Statistical Mechanics II," Phys. Rev. **108**, 171–190 (1957).

[26] E. T. Jaynes, "Information Theory and Statistical Mechanics (Notes by the lecturer)," in Statistical physics 3 (1963), pp. 181–218.

[27] E. T. Jaynes, *Probability Theory: The Logic of Science*, edited by G. L. Bretthorst (Cambridge University Press, 2003).

[28] P. J. Coles, M. Berta, M. Tomamichel, and S. Wehner, "Entropic uncertainty relations and their applications," Rev. Mod. Phys. **89**, 015002 (2017).

[29] A. Hertz and N. J. Cerf, "Continuous-variable entropic uncertainty relations," J. Phys. A Math. Theor. **52**, 173001 (2019).

[30] H. Maassen and J. B. M. Uffink, "Generalized entropic uncertainty relations," Phys. Rev. Lett. **60**, 1103–1106 (1988).

[31] M. Berta, M. Christandl, R. Colbeck, J. M. Renes, and R. Renner, "The uncertainty principle in the presence of quantum memory," Nat. Phys. **6**, 659–662 (2010).

[32] I. Białynicki-Birula and J. Mycielski, "Uncertainty relations for information entropy in wave mechanics," Commun. Math. Phys. **44**, 129–132 (1975).

[33] R. L. Frank and E. H. Lieb, "Entropy and the Uncertainty Principle," Ann. Henri Poincaré **13**, 1711–1717 (2012).

[34] E. H. Lieb, "Proof of an entropy conjecture of Wehrl," Commun. Math. Phys. **62**, 35–41 (1978).

[35] N. J. Cerf and C. Adami, "Negative Entropy and Information in Quantum Mechanics," Phys. Rev. Lett. **79**, 5194–5197 (1997).

[36] S. Popescu, A. Short, and A. Winter, "Entanglement and the foundations of statistical mechanics," Nat. Phys. **2**, 754–758 (2006).

[37] A. M. Kaufman, M. E. Tai, A. Lukin, M. Rispoli, R. Schittko, P. M. Preiss, and M. Greiner, "Quantum thermalization through entanglement in an isolated many-body system," Science **353**, 794–800 (2016).

[38] R. Islam, R. Ma, P. M. Preiss, M. E. Tai, A. Lukinand, M. Rispoli, and M. Greiner, "Measuring entanglement entropy in a quantum many-body system," Nature **528**, 77–83 (2015).

[39] J. Maldacena and L. Susskind, "Cool horizons for entangled black holes," Fortschr. Phys. **61**, 781–811 (2013).

[40] S. N. Solodukhin, "Entanglement Entropy of Black Holes," Living Rev. Relativ. **14**, 8 (2011).

[41] T. Nishioka, S. Ryu, and T. Takayanagi, "Holographic entanglement entropy: an overview," J. Phys. A Math. Theor. **42**, 504008 (2009).

[42] O. Gühne and G. Tóth, "Entanglement detection," Phys. Rep. **474**, 1–75 (2009).

[43] L.-M. Duan, G. Giedke, J. I. Cirac, and P. Zoller, "Inseparability Criterion for Continuous Variable Systems," Phys. Rev. Lett. **84**, 2722–2725 (2000).

[44] R. Simon, "Peres-Horodecki Separability Criterion for Continuous Variable Systems," Phys. Rev. Lett. **84**, 2726–2729 (2000).

[45] S. Mancini, V. Giovannetti, D. Vitali, and P. Tombesi, "Entangling Macroscopic Oscillators Exploiting Radiation Pressure," Phys. Rev. Lett. **88**, 120401 (2002).

[46] V. Giovannetti, S. Mancini, D. Vitali, and P. Tombesi, "Characterizing the entanglement of bipartite quantum systems," Phys. Rev. A **67**, 022320 (2003).

[47] S. P. Walborn, B. G. Taketani, A. Salles, F. Toscano, and R. L. de Matos Filho, "Entropic Entanglement Criteria for Continuous Variables," Phys. Rev. Lett. **103**, 160505 (2009).

[48] S. P. Walborn, A. Salles, R. M. Gomes, F. Toscano, and P. H. S. Ribeiro, "Revealing Hidden Einstein-Podolsky-Rosen Nonlocality," Phys. Rev. Lett. **106**, 130402 (2011).

[49] A. Saboia, F. Toscano, and S. P. Walborn, "Family of continuous-variable entanglement criteria using general entropy functions," Phys. Rev. A **83**, 032307 (2011).

[50] J. Schneeloch and G. A. Howland, "Quantifying high-dimensional entanglement with Einstein-Podolsky-Rosen correlations," Phys. Rev. A **97**, 042338 (2018).

[51] J. Schneeloch, C. C. Tison, M. L. Fanto, P. M. Alsing, and G. A. Howland, "Quantifying entanglement in a 68-billion-dimensional quantum state space," Nat. Commun. **10**, 1–7 (2019).

[52] H. Nyquist, "Certain Factors Affecting Telegraph Speed," Trans. AIEE **XLIII**, 412–422 (1924).

[53] H. Nyquist, "Certain Topics in Telegraph Transmission Theory," Trans. AIEE **47**, 617–644 (1928).

[54] R. V. L. Hartley, "Transmission of Information," Bell Syst. Tech. J. **7**, 535–563 (1928).

[55] C. E. Shannon and W. Weaver, *The Mathematical Theory of Communication* (The University Of Illinois Press, 1964).

[56] T. M. Cover and J. A. Thomas, *Elements of Information Theory, Second Edition* (John Wiley and Sons, 2006).

[57] P.-S. Laplace, *A Philosophical Essay On Probabilities*, trans. by F. W. Truscott and F. L. Emory (John Wiley & Sons and Chapman & Hall, 1902).

[58] A. Kolmogorov, *Grundbegriffe der Wahrscheinlichkeitsrechnung* (Verlag von Julius Springer, 1933).

[59] V. Vedral, "The role of relative entropy in quantum information theory," Rev. Mod. Phys. **74**, 197–234 (2002).

[60] M. Tribus and E. C. McIrvine, "Energy and Information," Sci. Am. **225**, 179–190 (1971).

[61] R. Jancel, *Foundations of Classical and Quantum Statistical Mechanics: International Series of Monographs in Natural Philosophy* (Elsevier, 1963).

[62] W. P. Schleich, *Quantum Optics in Phase Space* (Wiley-VCH Verlag Berlin, 2001).

[63] C. Weedbrook, S. Pirandola, R. García-Patrón, N. J. Cerf, T. C. Ralph, J. H. Shapiro, and S. Lloyd, "Gaussian quantum information," Rev. Mod. Phys. **84**, 621 (2012).

[64] A. Hertz, "Exploring continuous-variable entropic uncertainty relations and separability criteria in quantum phase space," PhD thesis (Université libre de Bruxelles, 2018).

[65] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, "Quantum entanglement," Rev. Mod. Phys. **81**, 865–942 (2009).

[66] J. Bub, "Quantum Mechanics is About Quantum Information," Found. Phys. **35**, 541–560 (2005).

[67] K. Kraus, *States, Effects and Operations: Fundamental Notions of Quantum Theory*, Lecture Notes in Physics 180 (Springer-Verlag Berlin Heidelberg, 1983).

[68] M. Ozawa, "Quantum measuring processes of continuous observables," J. Math. Phys. **25**, 79–87 (1984).

[69] V. Vedral, M. B. Plenio, M. A. Rippin, and P. L. Knight, "Quantifying Entanglement," Phys. Rev. Lett. **78**, 2275–2279 (1997).

[70] P. Pechukas, "Reduced Dynamics Need Not Be Completely Positive," Phys. Rev. Lett. **73**, 1060–1062 (1994).

[71] S. Milz, F. A. Pollock, and K. Modi, "An Introduction to Operational Quantum Dynamics," Open Syst. Inf. Dyn. **24**, 1740016 (2017).

[72] R. Beneduci, "Notes on Naimark's dilation theorem," J. Phys. Conf. Ser. **1638**, 012006 (2020).

[73] A. Mueller-Hermes and D. Reeb, "Monotonicity of the Quantum Relative Entropy Under Positive Maps," Ann. Henri Poincaré **18**, 1777–1788 (2017).

[74] M. Berta, M. Lemm, and M. M. Wilde, "Monotonicity of quantum relative entropy and recoverability," Quantum Inf. Comp. **15**, 1333–1354 (2015).

[75] A. A. Kuznetsova, "Conditional Entropy for Infinite-Dimensional Quantum Systems," Theory Probab. Its Appl. **55**, 709–717 (2011).

[76] S. A. S. Eddington, *The Nature of the Physical World* (Cambridge University Press, 1948).

[77] P. Walters, *An Introduction to Ergodic Theory* (Springer-Verlag New York, 1982).

[78] J. M. Deutsch, "Quantum statistical mechanics in a closed system," Phys. Rev. A **43**, 2046–2049 (1991).

[79] M. Srednicki, "Chaos and quantum thermalization," Phys. Rev. E **50**, 888–901 (1994).

[80] M. Rigol, V. Dunjko, and M. Olshanii, "Thermalization and its mechanism for generic isolated quantum systems," Nature **452**, 854–858 (2008).

[81] L. D'Alessio, Y. Kafri, A. Polkovnikov, and M. Rigol, "From quantum chaos and eigenstate thermalization to statistical mechanics and thermodynamics," Adv. Phys. **65**, 239–362 (2016).

[82] J. M. Deutsch, "Eigenstate thermalization hypothesis," Rep. Prog. Phys. **81**, 082001 (2018).

[83] E. H. Lieb and J. Yngvason, "The physics and mathematics of the second law of thermodynamics," Phys. Rep. **310**, 1–96 (1999).

[84] E. H. Lieb and J. Yngvason, "The entropy concept for non-equilibrium states," Proc. Phys. Soc. A **469**, 1–15 (2013).

[85] E. H. Lieb and J. Yngvason, "Entropy meters and the entropy of non-extensive systems," Proc. Phys. Soc. A **470**, 1–9 (2014).

[86] E. Zanchini, "On the definition of extensive property energy by the first postulate of thermodynamics," Found. Phys. **16**, 923–935 (1986).

[87] E. Zanchini, "Thermodynamics: Energy of closed and open systems," Il Nuovo Cimento B **101**, 453–465 (1988).

[88] E. Zanchini, "Thermodynamics: Energy of nonsimple systems and second postulate," Il Nuovo Cimento B **107**, 123–139 (1992).

[89] E. Zanchini and G. P. Beretta, "Rigorous Axiomatic Definition of Entropy Valid Also for Non-Equilibrium States," AIP Conf. Proc. **1033**, 296–301 (2008).

[90] E. Zanchini and G. P. Beretta, "Removing Heat and Conceptual Loops from the Definition of Entropy," Int. J. Thermodyn. **13**, 67–76 (2010).

[91] G. P. Beretta and E. Zanchini, "Rigorous and General Definition of Thermodynamic Entropy," in *Thermodynamics* (IntechOpen, 2011), pp. 23–50.

[92] E. Zanchini and G. P. Beretta, "Recent Progress in the Definition of Thermodynamic Entropy," Entropy **16**, 1547–1570 (2014).

[93] G. P. Beretta and E. Zanchini, "New definitions of thermodynamic temperature and entropy not based on the concepts of heat and thermal reservoir," Atti della Accademia Peloritana dei Pericolanti - Classe di Scienze Fisiche, Matematiche e Naturali **97**, 1–28 (2019).

[94] W.-M. Zhang, D. H. Feng, and R. Gilmore, "Coherent states: Theory and some applications," Rev. Mod. Phys. **62**, 867–927 (1990).

[95] A. M. Perelomov, "Coherent states for arbitrary Lie group," Commun. Math. Phys. **26**, 222–236 (1972).

[96] J. M. Radcliffe, "Some properties of coherent spin states," J. Phys. A **4**, 313–323 (1971).

[97] R. Gilmore, "On Properties Of Coherent States," Rev. Mex. de Fis. **23**, 143–187 (1974).

[98] A. M. Perelomov, *Generalized Coherent States and Their Applications* (Springer, Berlin, Heidelberg, 1986).

[99] A. Z. Goldberg, A. B. Klimov, M. Grassl, G. Leuchs, and L. L. Sánchez-Soto, "Extremal quantum states," AVS Quantum Sci. **2**, 044701 (2020).

[100] K. Husimi, "Some formal properties of the density matrix," Proc. Phys.-Math. Soc. Jap. 3rd Ser. **22**, 264–314 (1940).

[101] N. D. Cartwright, "A non-negative Wigner-type distribution," Phys. A **83**, 210–212 (1976).

[102] H.-W. Lee, "Theory and application of the quantum phase-space distribution functions," Phys. Rep. **259**, 147–211 (1995).

[103] E. P. Wigner, "On the Quantum Correction For Thermodynamic Equilibrium," Phys. Rev. **40**, 749–759 (1932).

[104] R. J. Glauber, "Coherent and Incoherent States of the Radiation Field," Phys. Rev. **131**, 2766–2788 (1963).

[105] E. C. G. Sudarshan, "Equivalence of Semiclassical and Quantum Mechanical Descriptions of Statistical Light Beams," Phys. Rev. Lett. **10**, 277–279 (1963).

[106] R. Hudson, "When is the wigner quasi-probability density non-negative?" Rep. Math. Phys. **6**, 249–252 (1974).

[107] J. Williamson, "On the Algebraic Problem Concerning the Normal Forms of Linear Dynamical Systems," Am. J. Math. **58**, 141–163 (1936).

[108] W. Heisenberg, "Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik," Z. Phys. **43**, 172–198 (1927).

[109] H. P. Robertson, "The Uncertainty Principle," Phys. Rev. **34**, 163–164 (1929).

[110] E. Schrödinger, "Zum Heisenbergschen Unschärfeprinzip," Sitzungsberichte der Preußischen Akademie der Wissenschaften. Physikalisch-mathematische Klasse **14**, 296–303 (1930).

[111] H. P. Robertson, "A general formulation of the uncertainty principle and its classical interpretation," Phys. Rev. **35**, 667 (1930).

[112] R. Simon, N. Mukunda, and B. Dutta, "Quantum-noise matrix for multimode systems: U(n) invariance, squeezing, and normal forms," Phys. Rev. A **49**, 1567–1583 (1994).

[113] I. Białynicki-Birula and Ł. Rudnicki, "Entropic Uncertainty Relations in Quantum Physics," in *Statistical Complexity* (Springer, Dordrecht, 2011).

[114] H. Everett, "'Relative State' Formulation of Quantum Mechanics," Rev. Mod. Phys. **29**, 454–462 (1957).

[115] I. I. Hirschman, "A Note on Entropy," Am. J. Math. **79**, 152–156 (1957).

[116] W. Beckner, "Inequalities in Fourier Analysis," Ann. Math. **102**, 159–182 (1975).

[117] E. H. Kennard, "Zur Quantenmechanik einfacher Bewegungstypen," Z. Phys. **44**, 326–352 (1927).

[118] H. Weyl, *Gruppentheorie und Quantenmechanik* (Hirzel, Leipzig, 1928).

[119] A. Hertz, M. G. Jabbour, and N. J. Cerf, "Entropy-power uncertainty relations: towards a tight inequality for all Gaussian pure states," J. Phys. A Math. Theor. **50**, 385301 (2017).

[120] A. Hertz, O. Oreshkov, and N. J. Cerf, "Multicopy uncertainty observable inducing a symplectic-invariant uncertainty relation in position and momentum phase space," Phys. Rev. A **100**, 052112 (2019).

[121] Z. Van Herstraeten and N. J. Cerf, "Quantum Wigner entropy," Phys. Rev. A **104**, 042211 (2021).

[122] Z. V. Herstraeten, M. G. Jabbour, and N. J. Cerf, "Continuous majorization in quantum phase space," arXiv:2108.09167, 1–13 (2021).

[123] M. Rumin, "Balanced distribution-energy inequalities and related entropy bounds," Duke Math. J. **160**, 567–597 (2011).

[124] E. H. Lieb and J. P. Solovej, "Proof of an entropy conjecture for Bloch coherent spin states and its generalizations," Acta Math. **212**, 379–398 (2014).

[125] M. Grabowski, "Wehrl-Lieb's inequality for entropy and the uncertainty relation," Rep. Math. Phys. **20**, 153–155 (1984).

[126] E. A. Carlen, "Some integral identities and inequalities for entire functions and their application to the coherent state transform," J. Funct. Anal. **97**, 231–249 (1991).

[127] E. H. Lieb and R. Seiringer, "Stronger subadditivity of entropy," Phys. Rev. A **71**, 062329 (2005).

[128] M. Ohya and D. Petz, *Quantum Entropy and Its Use* (Springer-Verlag Berlin Heidelberg, 1993).

[129] E. H. Lieb and J. P. Solovej, "Proof of the Wehrl-type Entropy Conjecture for Symmetric $SU(N)$ Coherent States," Commun. Math. Phys. **348**, 567–578 (2016).

[130] E. H. Lieb and J. P. Solovej, "Wehrl-type coherent state entropy inequalities for SU(1,1) and its AX+B subgroup," in *Partial differential equations, spectral theory, and mathematical physics* (2021), pp. 301–314.

[131] D. Deutsch, "Uncertainty in Quantum Measurements," Phys. Rev. Lett. **50**, 631–633 (1983).

[132] K. Kraus, "Complementary observables and uncertainty relations," Phys. Rev. D **35**, 3070–3075 (1987).

[133] P. J. Coles, R. Colbeck, L. Yu, and M. Zwolak, "Uncertainty Relations from Simple Entropic Properties," Phys. Rev. Lett. **108**, 210405 (2012).

[134] M. Krishna and K. R. Parthasarathy, "An Entropic Uncertainty Principle for Quantum Measurements," Ind. J. Stat. **64**, 842–851 (2002).

[135] M. Tomamichel, "A Framework for Non-Asymptotic Quantum Information Theory," PhD thesis (Eidgenössische Technische Hochschule Zürich, 2012).

[136] P. J. Coles and M. Piani, "Improved entropic uncertainty relations and information exclusion relations," Phys. Rev. A **89**, 022112 (2014).

[137] M. Rumin, "An Entropic Uncertainty Principle for Positive Operator Valued Measures," Lett. Math. Phys. **100**, 291–308 (2012).

[138] G. D. Palma and D. Trevisan, "The Conditional Entropy Power Inequality for Bosonic Quantum Systems," Commun. Math. Phys. **360**, 639–662 (2018).

[139] G. D. Palma, "The Wehrl entropy has Gaussian optimizers," Lett. Math. Phys. **108**, 97–116 (2018).

[140] J. M. Renes and J.-C. Boileau, "Conjectured Strong Complementary Information Tradeoff," Phys. Rev. Lett. **103**, 020402 (2009).

[141] M. Tomamichel and R. Renner, "Uncertainty Relation for Smooth Entropies," Phys. Rev. Lett. **106**, 110506 (2011).

[142] B. Bergh and M. Gärttner, "Entanglement detection in quantum many-body systems using entropic uncertainty relations," Phys. Rev. A **103**, 052412 (2021).

[143] B. Bergh and M. Gärttner, "Experimentally Accessible Bounds on Distillable Entanglement from Entropic Uncertainty Relations," Phys. Rev. Lett. **126**, 190503 (2021).

[144] R. L. Frank and E. H. Lieb, "Extended Quantum Conditional Entropy and Quantum Uncertainty Inequalities," Commun. Math. Phys. **323**, 487–495 (2013).

[145] F. Furrer, M. Berta, M. Tomamichel, V. B. Scholz, and M. Christandl, "Position-momentum uncertainty relations in the presence of quantum memory," J. Math. Phys. **55**, 122205 (2014).

[146] M. J. W. Hall, "Information Exclusion Principle for Complementary Observables," Phys. Rev. Lett. **74**, 3307–3311 (1995).

[147] M. J. W. Hall, "Quantum information and correlation bounds," Phys. Rev. A **55**, 100–113 (1997).

[148] P. Busch, P. Lahti, and R. F. Werner, "Colloquium: Quantum root-mean-square error and measurement uncertainty relations," Rev. Mod. Phys. 86, 1261–1281 (2014).

[149] F. Buscemi, M. J. W. Hall, M. Ozawa, and M. M. Wilde, "Noise and Disturbance in Quantum Measurements: An Information-Theoretic Approach," Phys. Rev. Lett. 112, 050401 (2014).

[150] M. J. W. Hall, "Almost-periodic time observables for bound quantum systems," J. Phys. A Math. Theor. 41, 255301 (2008).

[151] A. Boette, R. Rossignoli, N. Gigena, and M. Cerezo, "System-time entanglement in a discrete-time model," Phys. Rev. A 93, 062127 (2016).

[152] M. J. W. Hall, "Entropic Heisenberg limits and uncertainty relations from the Holevo information bound," J. Phys. A Math. Theor. 51, 364001 (2018).

[153] P. J. Coles, V. Katariya, S. Lloyd, I. Marvian, and M. M. Wilde, "Entropic Energy-Time Uncertainty Relation," Phys. Rev. Lett. 122, 100401 (2019).

[154] I. Bialynicki-Birula, "Formulation of the uncertainty relations in terms of the Rényi entropies," Phys. Rev. A 74, 052101 (2006).

[155] H. Casini, M. Huerta, J. M. Magán, and D. Pontello, "Entanglement entropy and superselection sectors. Part I. Global symmetries," J. High Energy Phys. 2020, 14 (2020).

[156] H. Casini, M. Huerta, J. M. Magán, and D. Pontello, "Entropic order parameters for the phases of QFT," J. High Energy Phys. 2021, 277 (2021).

[157] J. M. Magan and D. Pontello, "Quantum complementarity through entropic certainty principles," Phys. Rev. A 103, 012211 (2021).

[158] L. Susskind and A. Friedman, *Quantum Mechanics: The Theoretical Minimum* (Penguin, 2015).

[159] M. B. Plenio and S. Virmani, "An Introduction to Entanglement Measures," Quantum Inf. Comput. 7, 1–51 (2007).

[160] V. Vedral and M. B. Plenio, "Entanglement measures and purification procedures," Phys. Rev. A 57, 1619–1633 (1998).

[161] S. Gharibian, "Strong NP-Hardness of the Quantum Separability Problem," Quantum Inf. Comput. 10, 343–360 (2010).

[162] Y. Huang, "Computing quantum discord is NP-complete," 16, 033027 (2014).

[163] A. Peres, "Separability Criterion for Density Matrices," Phys. Rev. Lett. 77, 1413–1415 (1996).

[164] M. Horodecki, P. Horodecki, and R. Horodecki, "Separability of mixed states: necessary and sufficient conditions," Phys. Lett. A **223**, 1–8 (1996).

[165] M. Horodecki, P. Horodecki, and R. Horodecki, "Mixed-State Entanglement and Distillation: Is there a "Bound" Entanglement in Nature?" Phys. Rev. Lett. **80**, 5239–5242 (1998).

[166] H. Nha and M. S. Zubairy, "Uncertainty Inequalities as Entanglement Criteria for Negative Partial-Transpose States," Phys. Rev. Lett. **101**, 130402 (2008).

[167] A. Einstein, B. Podolsky, and N. Rosen, "Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?" Phys. Rev. **47**, 777–780 (1935).

[168] E. Shchukin and W. Vogel, "Inseparability Criteria for Continuous Bipartite Quantum States," Phys. Rev. Lett. **95**, 230502 (2005).

[169] E. V. Shchukin and W. Vogel, "Nonclassical moments and their measurement," Phys. Rev. A **72**, 043808 (2005).

[170] M. Horodecki and P. Horodecki, "Reduction criterion of separability and limits for a class of distillation protocols," Phys. Rev. A **59**, 4206–4216 (1999).

[171] N. J. Cerf, C. Adami, and R. M. Gingrich, "Reduction criterion for separability," Phys. Rev. A **60**, 898–909 (1999).

[172] O. Gühne, "Characterizing Entanglement via Uncertainty Relations," Phys. Rev. Lett. **92**, 117903 (2004).

[173] O. Gühne and M. Lewenstein, "Entropic uncertainty relations and entanglement," Phys. Rev. A **70**, 022316 (2004).

[174] V. Giovannetti, S. Lloyd, and L. Maccone, "Advances in quantum metrology," Nat. Photon. **5**, 222–229 (2011).

[175] L. Bombelli, R. K. Koul, J. Lee, and R. D. Sorkin, "Quantum source of entropy for black holes," Phys. Rev. D **34**, 373–383 (1986).

[176] M. Srednicki, "Entropy and area," Phys. Rev. Lett. **71**, 666–669 (1993).

[177] C. Callan and F. Wilczek, "On geometric entropy," Phys. Lett. B **333**, 55–61 (1994).

[178] J. H. Cooperman and M. A. Luty, "Renormalization of entanglement entropy and the gravitational effective action," J. High Energy Phys. **2014**, 45 (2014).

[179] S. Ryu and T. Takayanagi, "Aspects of holographic entanglement entropy," J. High Energy Phys. **2006**, 045 (2006).

[180] S. Ryu and T. Takayanagi, "Holographic Derivation of Entanglement Entropy from the AdS/CFT Correspondence," Phys. Rev. Lett. **96**, 181602 (2006).

[181] H. Casini, M. Huerta, and R. C. Myers, "Towards a derivation of holographic entanglement entropy," J. High Energy Phys. **2011**, 36 (2011).

[182] L. Amico, R. Fazio, A. Osterloh, and V. Vedral, "Entanglement in many-body systems," Rev. Mod. Phys. **80**, 517–576 (2008).

[183] J. Eisert, M. Cramer, and M. B. Plenio, "Colloquium: Area laws for the entanglement entropy," Rev. Mod. Phys. **82**, 277–306 (2010).

[184] P. Calabrese, J. Cardy, and B. Doyon, "Entanglement entropy in extended quantum systems," J. Phys. A Math. Theor. **42**, 500301 (2009).

[185] D. E. Kharzeev and E. M. Levin, "Deep inelastic scattering as a probe of entanglement," Phys. Rev. D **95**, 114008 (2017).

[186] E. Shuryak and I. Zahed, "Regimes of the Pomeron and its Intrinsic Entropy," Ann. Phys. (NY) **396**, 1–17 (2018).

[187] J. Berges, S. Floerchinger, and R. Venugopalan, "Dynamics of entanglement in expanding quantum fields," J. High Energy Phys. **2018**, 145 (2018).

[188] J. Berges, S. Floerchinger, and R. Venugopalan, "Thermal excitation spectrum from entanglement in an expanding quantum string," Phys. Lett. B **778**, 442–446 (2018).

[189] A. Kovner, M. Lublinsky, and M. Serino, "Entanglement entropy, entropy production and time evolution in high energy QCD," Phys. Lett. B **792**, 4–15 (2019).

[190] N. Armesto, F. Domínguez, A. Kovner, M. Lublinsky, and V. V. Skokov, "The Color Glass Condensate density matrix: Lindblad evolution, entanglement entropy and Wigner functional," J. High Energy Phys. **2019**, 25 (2019).

[191] Z. Tu, D. E. Kharzeev, and T. Ullrich, "The EPR paradox and quantum entanglement at sub-nucleonic scales," Phys. Rev. Lett. **124**, 062001 (2020).

[192] B. Hatfield, *Quantum Field Theory Of Point Particles And Strings* (CRC Press, 2018).

[193] E. Witten, "APS Medal for Exceptional Achievement in Research: Invited article on entanglement properties of quantum field theory," Rev. Mod. Phys. **90**, 045003 (2018).

[194] H. Casini and M. Huerta, "Entanglement entropy in free quantum field theory," J. Phys. A Math. Theor. **42**, 504007 (2009).

[195] P. Calabrese and J. Cardy, "Entanglement entropy and quantum field theory," J. Stat. Mech. Theory Exp. **2004**, P06002 (2004).

[196] M. Headrick, "Lectures on entanglement entropy in field theory and holography," arXiv:1907.08126, 1–66 (2019).

[197] K. Osterwalder and R. Schrader, "Axioms for Euclidean Green's functions," Commun. Math. Phys. **31**, 83–112 (1973).

[198] K. Osterwalder and R. Schrader, "Axioms for Euclidean Green's functions II," Commun. Math. Phys. **42**, 281–305 (1975).

[199] R. Haag, *Local quantum physics: Fields, particles, algebras* (Springer-Verlag Berlin Heidelberg, 1996).

[200] A. Jaffe and E. Witten, "Quantum Yang-Mills Theory," Clay Mathematics Institute, 1–14 (2000).

[201] L. V. Keldysh, "Diagram Technique For Nonequilibrium Processes," Soviet Physics JETP **20**, 1018–1026 (1964).

[202] L. P. Kadanoff and G. Baym, *Quantum Statistical Mechanics: Green's Function Methods in Equilibrium and Nonequilibrium Problems* (CRC Press, 1962).

[203] H. Reeh and S. Schlieder, "Bemerkungen zur unitäräquivalenz von lorentzinvarianten feldern," Il Nuovo Cimento (1955-1965) **22**, 1051–1068 (1961).

[204] D. N. Page, "Average entropy of a subsystem," Phys. Rev. Lett. **71**, 1291–1294 (1993).

[205] S. K. Foong and S. Kanno, "Proof of Page's conjecture on the average entropy of a subsystem," Phys. Rev. Lett. **72**, 1148–1151 (1994).

[206] S. Sen, "Average Entropy of a Quantum Subsystem," Phys. Rev. Lett. **77**, 1–3 (1996).

[207] T. Hartman, *Lectures on Quantum Gravity and Black Holes*, 2015.

[208] H. Araki, "Relative Entropy for States of von Neumann Algebras II," Publ. Res. Inst. Math. Sci. **13**, 173–192 (1977).

[209] S. Hollands and K. Sanders, *Entanglement Measures and their Properties in Quantum Field Theory* (Springer, 2018).

[210] H. Casini, "Relative entropy and the Bekenstein bound," Class. Quantum Gravity **25**, 205021 (2008).

[211] R. Bousso, H. Casini, Z. Fisher, and J. Maldacena, "Proof of a quantum Bousso bound," Phys. Rev. D **90**, 044002 (2014).

[212] R. Bousso, H. Casini, Z. Fisher, and J. Maldacena, "Entropy on a null surface for interacting quantum field theories and the Bousso bound," Phys. Rev. D **91**, 084030 (2015).

[213] D. D. Blanco, H. Casini, L.-Y. Hung, and R. C. Myers, "Relative entropy and holography," J. High Energy Phys. **2013**, 60 (2013).

[214] A. C. Wall, "Proof of the generalized second law for rapidly evolving Rindler horizons," Phys. Rev. D **82**, 124019 (2010).

[215] A. C. Wall, "Proof of the generalized second law for rapidly changing fields and arbitrary horizon slices," Phys. Rev. D **85**, 104049 (2012).

[216] A. C. Wall, "Erratum: Proof of the generalized second law for rapidly changing fields and arbitrary horizon slices [Phys. Rev. D 85, 104049 (2012)]," Phys. Rev. D **87**, 069904 (2013).

[217] R. E. Arias, D. D. Blanco, H. Casini, and M. Huerta, "Local temperatures and local terms in modular Hamiltonians," Phys. Rev. D **95**, 065005 (2017).

[218] N. Lashkari, "Relative Entropies in Conformal Field Theory," Phys. Rev. Lett. **113**, 051602 (2014).

[219] N. Lashkari, "Modular Hamiltonian for Excited States in Conformal Field Theory," Phys. Rev. Lett. **117**, 041601 (2016).

[220] G. Sárosi and T. Ugajin, "Relative entropy of excited states in two dimensional conformal field theories," J. High Energy Phys. **2016**, 114 (2016).

[221] T. Ugajin, "Mutual information of excited states and relative entropy of two disjoint subsystems in CFT," J. High Energy Phys. **2017**, 184 (2017).

[222] P. Ruggiero and P. Calabrese, "Relative entanglement entropies in 1+1-dimensional conformal field theories," J. High Energy Phys. **2017**, 39 (2017).

[223] D. L. Jafferis, A. Lewkowycz, J. Maldacena, and S. J. Suh, "Relative entropy equals bulk relative entropy," J. High Energy Phys. **2016**, 4 (2016).

[224] T. Takayanagi, T. Ugajin, and K. Umemoto, "Towards an entanglement measure for mixed states in CFTs based on relative entropy," J. High Energy Phys. **2018**, 166 (2018).

[225] M. M. Wolf, "Violation of the Entropic Area Law for Fermions," Phys. Rev. Lett. **96**, 010404.

[226] D. Gioev and I. Klich, "Entanglement Entropy of Fermions in Any Dimension and the Widom Conjecture," Phys. Rev. Lett. **96**, 100503 (2006).

[227] E. Fradkin and J. E. Moore, "Entanglement Entropy of 2D Conformal Quantum Critical Points: Hearing the Shape of a Quantum Drum," Phys. Rev. Lett. **97**, 050404 (2006).

[228] R. Y. Rubinstein and D. P. Kroese, *The Cross-Entropy Method, A Unified Approach to Combinatorial Optimization, Monte-Carlo Simulation and Machine Learning* (Springer New York, 2004).

[229] P.-T. de Boer, D. P. Kroese, S. Mannor, and R. Y. Rubinstein, "A Tutorial on the Cross-Entropy Method," Ann. Oper. Res. **134**, 19–67 (2005).

[230] H. Jeffreys, "An Invariant Form for the Prior Probability in Estimation Problems," Proc. Phys. Soc. A **186**, 453–461 (1946).

[231] E. T. Jaynes, "Prior Probabilities," IEEE Trans. Syst. Cybern. **4**, 227–241 (1968).

[232] R. A. Fisher, "Theory of Statistical Estimation," Math. Proc. Camb. Philos. Soc. **22**, 700–725 (1925).

[233] S.-i. Amari, *Information Geometry and Its Applications* (Springer, Tokyo, 2016).

[234] M. L. Gromov, *In a Search for a Structure, Part 1: On Entropy*, 2012.

[235] E. Hellinger, "Neue Begründung der Theorie quadratischer Formen von unendlichvielen Veränderlichen.," J. Reine Angew. Math. **1909**, 210–271 (1909).

[236] F. Nielsen, "An elementary introduction to information geometry," arXiv:1808.08271, 1–56 (2020).

[237] D. Petz and C. Ghinea, "Introduction to Quantum Fisher Information," in *Quantum probability and related topics* (2011), pp. 261–281.

[238] J. Liu, H. Yuan, X.-M. Lu, and X. Wang, "Quantum Fisher information matrix and multiparameter estimation," J. Phys. A Math. Theor. **53**, 023001 (2019).

[239] S. L. Braunstein and C. M. Caves, "Statistical distance and the geometry of quantum states," Phys. Rev. Lett. **72**, 3439–3443 (1994).

[240] P. Facchi, R. Kulkarnid, V. I. Man'ko, G. Marmo, E. C. G. Sudarshanh, and F. Ventrigliaf, "Classical and quantum fisher information in the geometrical formulation of quantum mechanics," Phys. Lett. A **374**, 4801–4803 (2010).

[241] H. Strobel, W. Muessel, D. Linnemann, T. Zibold, D. B. Hume, L. Pezzè, A. Smerzi, and M. K. Oberthaler, "Fisher information and entanglement of non-Gaussian spin states," Science **345**, 424–427 (2014).

[242] N. Lashkari and M. V. Raamsdonk, "Canonical energy is quantum Fisher information," J. High Energy Phys. **2016**, 153 (2016).

[243] P. Hauke, M. Heyl, L. Tagliacozzo, and P. Zoller, "Measuring multipartite entanglement through dynamic susceptibilities," Nat. Phys. **12**, 778–782 (2016).

[244] V. I. Man'ko, G. Marmo, F. Ventriglia, and P. Vitale, "Metric on the space of quantum states from relative entropy. Tomographic reconstruction," J. Phys. A Math. Theor. **50**, 335302 (2017).

[245] S. Banerjee, J. Erdmenger, and D. Sarkar, "Connecting Fisher information to bulk entanglement in holography," J. High Energy Phys., 1 (2018).

[246] D. Chruściński, F. M. Ciaglia, A. Ibort, G. Marmo, and F. Ventriglia, "Stratified manifold of quantum states, actions of the complex special linear group," Ann. Phys. (NY) **400**, 221–245 (2019).

[247] T. Sagawa, "Second Law-Like Inequalities with Quantum Relative Entropy: An Introduction," in *Lectures on quantum computing, thermodynamics and statistical physics* (2012), pp. 125–190.

[248] N. Dowling, "Formulating a Local Second Law of Thermodynamics Through Relative Entropy," Master Thesis (Heidelberg University, 2020).

[249] A. Einstein, "Theorie der Opaleszenz von homogenen Flüssigkeiten und Flüssigkeitsgemischen in der Nähe des kritischen Zustandes," Ann. Phys. **338**, 1275–1298 (1910).

[250] L. D. Landau and J. M. Lifschitz, "Vol. 5 - Statistical Physics, Part I," in *Course of Theoretical Physics*, 3rd (Pergamon Press Ltd., 1980).

[251] Y. Mishin, "Thermodynamic theory of equilibrium fluctuations," Ann. Phys. (NY) **363**, 48–97 (2015).

[252] G. N. Hatsopoulos and E. P. Gyftopoulos, "A unified quantum theory of mechanics and thermodynamics. Part I. Postulates," Found. Phys. **6**, 15–31 (1976).

[253] G. N. Hatsopoulos and E. P. Gyftopoulos, "A unified quantum theory of mechanics and thermodynamics. Part IIa. Available energy," Found. Phys. **6**, 127–141 (1976).

[254] G. N. Hatsopoulos and E. P. Gyftopoulos, "A unified quantum theory of mechanics and thermodynamics. Part IIb. Stable equilibrium states," Found. Phys. **6**, 439–455 (1976).

[255] G. N. Hatsopoulos and E. P. Gyftopoulos, "A unified quantum theory of mechanics and thermodynamics. Part III. Irreducible quantal dispersions," Found. Phys. **6**, 561–570 (1976).

[256] B. C. Eu, "Boltzmann entropy, relative entropy, and related quantities in thermodynamic space," J. Chem. Phys. **102**, 7169 (1995).

[257] P. Fraundorf, "Thermal roots of correlation-based complexity," Complexity **13**, 18–26 (2008).

[258] U. Heinz and R. Snellings, "Collective flow and viscosity in relativistic heavy-ion collisions," Annu. Rev. Nucl. Part. Sci. **63**, 123–151 (2013).

[259] W. Busza, K. Rajagopal, and W. van der Schee, "Heavy Ion Collisions: The Big Picture, and the Big Questions," Annu. Rev. Nucl. Part. Sci. **68**, 339–376 (2018).

[260] E. W. Kolb and M. S. Turner, *The Early Universe* (Frontiers in Physics, Vol. 69, 1990).

[261] S. Weinberg, *Cosmology* (Oxford University Press, 2008).

[262] T. Hatano and S.-i. Sasa, "Steady-State Thermodynamics of Langevin Systems," Phys. Rev. Lett. **86**, 3463–3466 (2001).

[263] H.-P. Breuer and F. Petruccione, *The theory of open quantum systems* (Oxford Scholarship Online, 2010).

[264] P. Kovtun, "Lectures on hydrodynamic fluctuations in relativistic theories," J. Phys. A Math. Theor. **45**, 473001 (2012).

[265] D. A. Teaney, "Viscous Hydrodynamics and the Quark Gluon Plasma," in *Quark-gluon plasma 4* (2010), pp. 207–266.

[266] S. Sachdev, *Quantum Phase Transitions*, 2nd ed. (Cambridge University Press, 2011).

[267] S. Floerchinger, "Variational principle for theories with dissipation from analytic continuation," J. High Energy Phys. **9**, 099 (2016).

[268] W. Israel and J. M. Stuart, "Transient relativistic thermodynamics and kinetic theory," Ann. Phys. (NY) **118**, 341–372 (1979).

[269] L. D. Landau and J. M. Lifschitz, "Vol. 6 - Fluid Mechanics," in *Course of Theoretical Physics*, 2nd (Pergamon Press Ltd., 1987).

[270] D. N. Zubarev, A. V. Prozorkevich, and S. A. Smolyanskii, "Derivation of nonlinear generalized equations of quantum relativistic hydrodynamics," Theor. Math. Phys. **40**, 821–831 (1979).

[271] F. Becattini, M. Buzzegoli, and E. Grossi, "Reworking Zubarev's approach to non-equilibrium quantum statistical mechanics," Particles **2**, 197–207 (2019).

[272] S. Floerchinger, T. Haas, and M. Schröfl, *Relative entanglement entropy of Gaussian models for scalar quantum field*, Work in progress, 2022.

[273] S. Floerchinger, T. Haas, and M. Schröfl, *Local Schwinger-Keldysh formalism*, Work in progress, 2022.

[274] R. Wald, *General Relativity* (The University of Chicago Press, 1984).

[275] E. Gourgoulhon, *Special Relativity in General Frames* (Springer, 2013).

[276] J. Lee, *Introduction to Smooth Manifolds* (Springer, 2012).

[277] F. Becattini and D. Rindori, "Extensivity, entropy current, area law and Unruh effect," Phys. Rev. D **99**, 125011 (2019).

[278] B. Höber, "An uncertainty relation for relative entropy," Bachelor Thesis (Heidelberg University, 2020).

[279] P. J. Coles, L. Yu, V. Gheorghiu, and R. B. Griffiths, "Information-theoretic treatment of tripartite systems and quantum channels," Phys. Rev. A 83, 062338 (2011).

[280] A. Barchielli, M. Gregoratti, and A. Toigo, "Measurement Uncertainty Relations for Position and Momentum: Relative Entropy Formulation," Entropy 19, 301 (2017).

[281] A. Barchielli, M. Gregoratti, and A. Toigo, "Measurement Uncertainty Relations for Discrete Observables: Relative Entropy Formulation," Commun. Math. Phys. 357, 1253–1304 (2018).

[282] W. Son, "Role of quantum non-Gaussian distance in entropic uncertainty relations," Phys. Rev. A 92, 012114 (2015).

[283] M. G. Genoni, M. G. A. Paris, and K. Banaszek, "Quantifying the non-Gaussian character of a quantum state by quantum relative entropy," Phys. Rev. A 78, 060303 (2008).

[284] S. M. Barnett and D. T. Pegg, "Quantum theory of rotation angles," Phys. Rev. A 41, 3427–3435 (1990).

[285] K. V. Mardia and P. E. Jupp, *Directional Statistics* (Wiley, 2009).

[286] M. Born, "Zur Quantenmechanik der Stoßvorgänge," Z. Phys. 37, 863–867 (1926).

[287] R. F. Streater and A. S. Wightman, *PCT, Spin and Statistics, and All That* (Princeton University Press, 2001).

[288] A. Vikman, "Suppressing Quantum Fluctuations in Classicalization," EPL 101, 34001 (2013).

[289] H. Müller-Groeling, "Wehrl entropy and Uncertainty Relations," Bachelor Thesis (Heidelberg University, 2020).

[290] R. J. Yáñez, W. V. Assche, and J. S. Dehesa, "Position and momentum information entropies of the D-dimensional harmonic oscillator and hydrogen atom," Phys. Rev. A 50, 3065–3079 (1994).

[291] V. Majerník and T. Opatrný, "Entropic uncertainty relations for a quantum oscillator," J. Phys. A Math. Theor. 29, 2187–2197 (1996).

[292] J. Sánchez-Ruiz, "Logarithmic potential of Hermite polynomials and information entropies of the harmonic oscillator eigenstates," J. Math. Phys. 38, 5031–5043 (1997).

[293] I. Toranzo and J. Dehesa, "Exact Shannon entropies for the multidimensional harmonic states," Phys. A 516, 273–279 (2019).

[294] A. Orłowski, "Wehrl's entropy and classification of states," Rep. Math. Phys. 43, 283–289 (1999).

[295] F. G. Mehler, "Ueber die Entwicklung einer Function von beliebig vielen Variablen nach Laplaceschen Functionen hoeherer Ordnung," J. Reine Angew. Math. **66**, 161–176 (1866).

[296] F. Mintert and K. Życzkowski, "Wehrl entropy, Lieb conjecture, and entanglement monotones," Phys. Rev. A **69**, 022317 (2004).

[297] R. Uola, T. Moroder, and O. Gühne, "Joint Measurability of Generalized Measurements Implies Classicality," Phys. Rev. Lett. **113**, 160403 (2014).

[298] R. Uola, C. Budroni, O. Gühne, and J.-P. Pellonpää, "One-to-One Mapping between Steering and Joint Measurability Problems," Phys. Rev. Lett. **115**, 230402 (2015).

[299] O. Gühne, E. Haapasalo, T. Kraft, J.-P. Pellonpää, and R. Uola, "Incompatible measurements in quantum information science," arXiv:2112.06784, 1–26 (2021).

[300] G. D. Palma, "Uncertainty relations with quantum memory for the Wehrl entropy," Commun. Math. Phys. **108**, 2139–2152 (2017).

[301] J. Schneeloch, C. J. Broadbent, and J. C. Howell, "Uncertainty relation for mutual information," Phys. Rev. A **90**, 062119 (2014).

[302] A. Serafini, *Quantum Continuous Variables, A Primer of Theoretical Methods* (CRC Press, 2017).

[303] L. Lami, A. Serafini, and G. Adesso, "Gaussian entanglement revisited," New J. Phys. **20**, 023030 (2018).

[304] A. Serafini, F. Illuminati, and S. D. Siena, "Symplectic invariants, entropic measures and correlations of Gaussian states," J. Phys. B **37**, L21–L28 (2003).

[305] S. Verdu and D. Guo, "A simple proof of the entropy-power inequality," IEEE Trans. Inf. Theory **52**, 2165–2166 (2006).

[306] O. Landon-Cardinal, L. C. G. Govia, and A. A. Clerk, "Quantitative Tomography for Continuous Variable Quantum Systems," Phys. Rev. Lett. **120**, 090501 (2018).

[307] J. W. Noh, A. Fougères, and L. Mandel, "Measurement of the quantum phase by photon counting," Phys. Rev. Lett. **67**, 1426–1429 (1991).

[308] J. W. Noh, A. Fougères, and L. Mandel, "Operational approach to the phase of a quantum field," Phys. Rev. A **45**, 424–442 (1992).

[309] U. Leonhardt and H. Paul, "Phase measurement and Q function," Phys. Rev. A **47**, R2460–R2463 (1993).

[310] P. Kunkel, M. Prüfer, S. Lannig, R. Rosa-Medina, A. Bonnin, M. Gärttner, H. Strobel, and M. K. Oberthaler, "Simultaneous Readout of Noncommuting Collective Spin Observables beyond the Standard Quantum Limit," Phys. Rev. Lett. **123**, 063603 (2019).

[311] P. Kunkel, M. Prüfer, S. Lannig, R. Strohmaier, M. Gärttner, H. Strobel, and M. K. Oberthaler, "Detecting Entanglement Structure in Continuous Many-Body Quantum Systems," Phys. Rev. Lett. **128**, 020402 (2022).

[312] F. Haas, J. Volz, R. Gehr, J. Reichel, and J. Estève, "Entangled states of more than 40 atoms in an optical fiber cavity," Science **344**, 180–183 (2014).

[313] G. Barontini, L. Hohmann, F. Haas, J. Estève, and J. Reichel, "Deterministic generation of multiparticle entanglement by quantum Zeno dynamics," Science **349**, 1317–1321 (2015).

[314] G. Kirchmair, B. Vlastakis, Z. Leghtas, S. E. Nigg, H. Paik, E. Ginossar, M. Mirrahimi, L. Frunzio, S. M. Girvin, and R. J. Schoelkopf, "Observation of quantum state collapse and revival due to the single-photon Kerr effect," Nature **495**, 205–209 (2013).

[315] S. Ditsch, "Wehrl Entropy and Quantum Fields," Bachelor Thesis (Heidelberg University, 2021).

[316] F. S. Schwark, "Entropic entanglement witness for the SU(2) phase space," Bachelor Thesis (Heidelberg University, 2022).

ACKNOWLEDGMENTS