

# Relative entropic uncertainty relation

Stefan Floerchinger<sup>\*,†</sup>, Tobias Haas<sup>‡</sup>, and Ben Hoerber<sup>‡</sup>

*Institut für Theoretische Physik, Universität Heidelberg, Philosophenweg 16, 69120 Heidelberg, Germany*



(Received 21 December 2020; accepted 25 May 2021; published 4 June 2021)

Quantum uncertainty relations are formulated in terms of relative entropy between distributions of measurement outcomes and suitable reference distributions with maximum entropy. This type of entropic uncertainty relation can be applied directly to observables with either discrete or continuous spectra. We find that a sum of relative entropies is bounded from above in a nontrivial way, which we illustrate with some examples.

DOI: [10.1103/PhysRevA.103.062209](https://doi.org/10.1103/PhysRevA.103.062209)

## I. INTRODUCTION

The uncertainty relation lies at the heart of quantum theory and is also well known among laymen. It was introduced first by Heisenberg in 1927 [1] and proven in the same year by Kennard [2] (see also [3]). It states the simple but astonishing fact that position  $x$  and momentum  $k$  cannot be fixed or measured with perfect precision at the same time. This idea was generalized to arbitrary Hermitian operators  $A$  and  $B$ ,

$$\sigma_A \sigma_B \geq \frac{1}{2} |\langle [A, B] \rangle|, \quad (1)$$

where  $\sigma_A$  denotes the standard deviation of  $A$ . Stated in this form, it becomes evident how the noncommutative nature of quantum mechanics is related to the uncertainty relation [4,5]. Thus another way of interpreting uncertainty relations is that the order of subsequent measurements matters.

Although the uncertainty relation formulated in terms of standard deviations is often handy, it is nowadays believed that a formulation in terms of *information entropies* is more adequate. One reason is that the standard deviation exhibits counterintuitive behavior in some situations [6]. Furthermore, entropy *is* a measure for the total uncertainty of a probability distribution [7]. A detailed discussion about advantages of uncertainty relations using entropy instead of variances can be found in Refs. [6,8–11].

The first entropic uncertainty relation (EUR) was formulated for the continuous variables position  $x$  and conjugate momentum  $k$ . In terms of the corresponding probability densities  $f(x)$  and  $g(k)$  it reads [10,12–15]

$$S(f) + S(g) \geq 1 + \ln \pi, \quad (2)$$

where  $S(f)$  denotes the differential entropy of the probability density  $f(x)$ . This entropic uncertainty relation is not only interesting because entropy is a more adequate measure of uncertainty, but also because of the implied position-momentum pair; hence it is stronger than Heisenberg's inequality. More precisely, it requires the distributions  $f(x)$  and  $g(k)$  to be of

Gaussian form in order to saturate the bound, which can be seen by rewriting it in terms of variances again [10,16].

Later, an entropic uncertainty relation for (nondegenerate) observables  $X$  and  $Z$  with *discrete* spectra

$$\begin{aligned} X|x\rangle &= x|x\rangle, \\ Z|z\rangle &= z|z\rangle \end{aligned} \quad (3)$$

was found by Deutsch [17] and improved by Maassen and Uffink [18] following a conjecture by Kraus [19]. Later, it was further strengthened by Berta *et al.* [20] (see also [21]), resulting in

$$S(p) + S(q) \geq \ln \frac{1}{c} + S(\rho). \quad (4)$$

Therein  $S(p)$  and  $S(q)$  denote the classical Shannon entropies [7] of the discrete distributions  $p(x)$  and  $q(z)$ , while the von Neumann entropy  $S(\rho)$  accounts for the *mixedness* of the system's state  $\rho$ . Moreover,  $c$  is defined as the maximum overlap between any two eigenvectors of  $X$  and  $Z$

$$c = \max_{x,z} |\langle x|z \rangle|^2, \quad (5)$$

which quantifies what one may call the *quantum incompatibility* of the two bases  $\{|x\rangle\}_x$  and  $\{|z\rangle\}_z$ .

The latter relation can also be generalized to two positive operator-valued probability measures (POVMs)  $\{\Lambda_x\}_x$  and  $\{\Gamma_z\}_z$ , which are non-negative operators that satisfy a completeness relation. Then, the Shannon entropies of the classical distributions  $p(x) = \text{Tr}\{\rho \Lambda_x\}$  and  $q(z) = \text{Tr}\{\rho \Gamma_z\}$  are bounded from below as prescribed by the Maassen-Uffink relation (4), if the quantum incompatibility gets replaced by [22–25]

$$c = \max_{x,z} \|\sqrt{\Lambda_x} \Gamma_z \sqrt{\Lambda_x}\|_\infty = \max_{x,z} \|\sqrt{\Lambda_x} \sqrt{\Gamma_z}\|_\infty^2, \quad (6)$$

where  $\|\cdot\|_\infty$  denotes the infinity operator norm. Note that, for projective measurements, this operator norm reduces to the overlap again.

Until today many different types of EURs were investigated, for example, in the context of Wehrl entropy [26], to classify the uncertainty between energy and time [27–30] or for the family of Rényi entropies [18,31]. Moreover, the idea of uncertainty relations was also extended to bi- and

\*stefan.floerchinger@thphys.uni-heidelberg.de

†t.haas@thphys.uni-heidelberg.de

‡hoeber@thphys.uni-heidelberg.de

tripartite systems with quantum memory in terms of conditional entropies [20,32–35]. Also, strengthened versions of the bound were investigated, which further reduce the quantum incompatibility measure  $c$  defined in (6) [23,24].

The two different types of observables were first unified in a formal way by Frank and Lieb [21] (using the Golden-Thompson inequality and the Gibbs variational principle) based on previous work by Rumin [36], who also generalized the relation to POVMs [37]. For the probability densities  $f(x)$  and  $g(k)$  their EUR reads

$$S(f) + S(g) \geq \ln(2\pi) + S(\rho), \quad (7)$$

while for discrete variables it reduces to the result of Maassen and Uffink (4). Interestingly, it is tighter than (2) for sufficiently mixed states and becomes tight in the infinite temperature limit  $\beta \rightarrow 0$ .

One of the main aims of this work is to present a simple EUR, which directly unifies the two cases of discrete and continuous random variables. In particular, we investigate a formulation in terms of *relative entropy* instead of entropy, which allows one to unify the two different types of observables in a straightforward manner.

Classically, the relative entropy (often also called Kullback-Leibler divergence) between two discrete distributions  $p(x)$  and  $\tilde{p}(x)$  is defined as [25,38–40]

$$S(p\|\tilde{p}) = \sum_x p(x) \ln[p(x)/\tilde{p}(x)] \quad (8)$$

and can be seen as a measure of their *distinguishability*. It cannot be considered as an actual distance measure since it does not obey a triangle inequality and is not symmetric. In fact it is a *divergence* in the mathematical sense.

Relative entropy plays a crucial role in the context of EURs. Some modern proofs of EURs for discrete variables [e.g., for the Maassen-Uffink relation (4)] rely on the monotonicity property of the quantum version of the relative entropy under completely positive trace-preserving maps<sup>1</sup> [25,34,41]. Furthermore, EURs can also be stated as information exclusion principles in the presence of classical or quantum memory [24,42,43]. In this sense, uncertainty between noncommuting observables is characterized by the impossibility of having a memory system that is strongly correlated to the measurements of both observables. Since mutual information can be expressed through relative entropy, such formulations can be regarded as EURs in terms of relative entropy in the presence of memory.

Moreover, the concept of relative entropy was used in Refs. [44,45] to formulate *measurement* EURs which deal with *joint approximate* measurements of incompatible observables (cf. Ref. [46] for an overview; see also [9]). In this field of research, EURs are used to characterize the quality of the aforementioned approximation or to constrain the disturbance of an observable after measuring another one (see, e.g., Ref. [47]), rather than to quantify the fundamental incompatibility of noncommuting observables, which is known as *preparation* uncertainty instead. In this work we only deal with the latter type of uncertainty.

In contrast to these considerations, our aim is to reformulate existing EURs in terms of relative entropies. The main result of this work is the relative entropic uncertainty relation (REUR) (54), which provides a bound for a sum of two relative entropies from above. Besides adding another perspective on EURs, the REUR is capable of describing the two cases of discrete and continuous variables simultaneously in one and the same inequality. This is achieved by picking suitable model distributions of maximum entropy  $p_{\max}$  and  $q_{\max}$ . In the case of discrete variables, convenient choices would be the uniform or a generalized Boltzmann distribution. For continuous variables, the most suitable choice is the Gaussian distribution, which also allows for a strengthened reformulation of Eq. (1) highlighting the looseness of the latter for non-Gaussian distributions. In this sense, the REUR (54) has to be understood as a reformulation of the Frank and Lieb EUR [21] in terms of relative entropies allowing for a more direct unification of the two cases of variables. While this formulation allows for a somewhat different physical interpretation, mathematically, it is equivalent to the relation proven in Ref. [21].

The paper is organized as follows. We begin with introducing the relative entropy in more detail in Sec. II. Then, we discuss the discrete case explicitly and show how to obtain an EUR in terms of relative entropy starting from Boltzmann distributions as optimal models in Sec. III. Afterwards, the continuous case is discussed in Sec. IV, where convenient model distributions turn out to be Gaussians. In Sec. V, we consider the general case and unify the relations obtained before. Furthermore, we make some remarks concerning the interpretation of the resulting EUR in terms of relative entropies. Also, we discuss a further nontrivial example, where we explicitly show that both sides of the REUR behave well when taking the continuum limit. Finally, we summarize our results and give an outlook in Sec. VI.

*Notation.* In this paper we adopt natural units, with  $\hbar = 1$ . Operator hats are dropped completely. Instead, we use capital letters for operators  $X$  and small letters for their eigenvalues  $x$  and eigenvectors  $|x\rangle$ . As a consequence, random variables are denoted by small letters, too. Furthermore, we use the symbol  $S$  for all kinds of entropies and take the underlying probability distribution  $p(x)$ , probability density  $f(x)$ , or density matrix  $\rho$  as argument. In this sense it depends on the argument whether  $S$  is a Shannon entropy  $S(p) = -\sum_x p(x) \ln p(x)$ , a differential entropy  $S(f) = -\int dx f(x) \ln f(x)$ , or a von Neumann entropy  $S(\rho) = -\text{Tr}\{\rho \ln \rho\}$ . It should be clear from the context what is meant in each case.

## II. RELATIVE ENTROPY

To get an intuitive understanding of relative entropy consider a random experiment where discrete events are correctly described by a distribution  $p(x)$ . If we mistakenly take the events to be distributed according to a model distribution  $\tilde{p}(x)$ , the relative entropy  $S(p\|\tilde{p})$  as defined in Eq. (8) quantifies the uncertainty deficit about  $p(x)$  due to the wrong model  $\tilde{p}(x)$ , more precisely the difference between the cross entropy  $S(p, \tilde{p}) = -\sum_x p(x) \ln \tilde{p}(x)$ , which is the expectation value of the apparent information content, and the real information content  $S(p) = -\sum_x p(x) \ln p(x)$  [48].

<sup>1</sup>This property is also often called *data-processing inequality*.

Consequently, relative entropy is a non-negative quantity being zero if and only if the two distributions agree. Moreover, to have a finite value, it needs the support condition  $\text{supp}[p(x)] \subseteq \text{supp}[\tilde{p}(x)]$  to hold; otherwise, its value is set to  $+\infty$ . The latter corresponds to the case where the model  $\tilde{p}(x)$  predicts zero probability for events, which can in fact happen. In this case one can rule out the model  $\tilde{p}(x)$  with certainty as soon as such an event occurs which causes the “distinguishability”  $S(p||\tilde{p})$  to be infinitely large. Thus a suitable model  $\tilde{p}(x)$  should at least cover all possible outcomes; otherwise, it could be ruled out for particular events [40]. Exchanging here  $p(x)$  and  $\tilde{p}(x)$  gives an intuitive understanding for the asymmetry of  $S(p||\tilde{p})$  under an exchange of its arguments.

It should be noted that the roles of the *true* and the *model* distribution can depend on the context. For our purposes it suffices to consider the second argument of relative entropy as the *model*.

Interestingly, relative entropy has some crucial advantages over entropy. Most importantly it turns out to be well defined for discrete and continuous random variables. Simply taking the continuum limit  $p(x) \rightarrow f(x)dx$  and  $\tilde{p}(x) \rightarrow \tilde{f}(x)dx$  yields [40]

$$S(f||\tilde{f}) = \int dx f(x) \ln[f(x)/\tilde{f}(x)]. \quad (9)$$

In contrast, Shannon’s entropy does not yield the differential entropy when taking this continuum limit [49]. As a consequence, differential entropy can become negative, while Shannon’s entropy is a non-negative quantity.

Moreover, relative entropy is invariant under a diffeomorphism  $x \rightarrow x'(x)$  on the underlying statistical manifold. This is also not the case for the differential entropy, which can be seen easily by considering scaling transformations. Thus it seems to be more appropriate to work with relative entropy instead of entropy, if one wants to describe discrete as well as continuous random variables.

### III. DISCRETE VARIABLES

We now consider first the case of discrete variables explicitly. We will show that suitable models are Boltzmann distributions, which directly leads us to a reformulation of the Maassen-Uffink inequality (4) in terms of relative entropies. Throughout the following section we assume the dimension of the Hilbert space to be finite  $d = \dim \mathcal{H} < \infty$ .

#### A. Quantum-classical states and classical distributions

We consider a quantum system in some arbitrary quantum state, which is specified by a non-negative, trace-class operator  $\rho$  with normalization  $\text{Tr}\{\rho\} = 1$ . We are interested in measuring two observables represented by Hermitian trace-class operators  $X$  and  $Z$ . Since we explicitly want to describe the discrete case, we assume that the spectra of both operators are bounded and discrete, such that the set of equations (3) holds. Therefore,  $\{|x\rangle\}_x$  and  $\{|z\rangle\}_z$  are orthonormal bases with normalizable elements.

The results of a measurement of  $X$  are then distributed according to

$$p(x) = \langle x|\rho|x\rangle, \quad (10)$$

i.e.,  $p(x)$  is the probability to obtain the value  $x$  when measuring  $X$  in its eigenbasis. The described procedure is the *projective measurement* (PVM), which is a special case of a POVM. In most generality, the probability distribution (10) can be constructed from the expectation value

$$p(x) = \text{Tr}\{\Lambda_x \rho\}, \quad (11)$$

which reduces to Eq. (10) if the POVM elements are chosen to be  $\Lambda_x = |x\rangle\langle x|$ , i.e., rank-1 projectors projecting the state  $\rho$  onto the  $x$ th eigenvalue of  $X$ .

In this setup, entropic uncertainty relations state that the sum of two classical entropies of the corresponding measured distributions  $p(x)$  and  $q(z)$  are bounded from below in a non-trivial way.

Furthermore, one can define a quantum-classical state  $\rho_X$  representing the state of the quantum system after measuring the operator  $X$  without recording the result. More precisely, the projective measurement is mathematically represented by applying a measurement map  $\mathcal{X}$  to the general density matrix  $\rho$  as follows:

$$\rho \rightarrow \rho_X = \mathcal{X}(\rho) = \sum_x p(x) |x\rangle\langle x|. \quad (12)$$

Therein  $p(x)$  is defined as in (10), such that  $\rho_X$  can indeed be interpreted as the classical state obtained after measuring the full quantum state  $\rho$  in the eigenbasis  $\{|x\rangle\}_x$  without recording the result. Consequently, the von Neumann entropy of a measured density operator  $\rho_X$  is the Shannon entropy of the classical probability distribution  $p(x)$

$$S(\rho_X) = - \sum_x p(x) \ln p(x) = S(p). \quad (13)$$

Since projective measurements are unital quantum channels, we have the following relation to the full quantum entropy [50]:

$$S(\rho) \leq S(p). \quad (14)$$

During a measurement without recording the result, the von Neumann entropy can increase but not decrease.

#### B. Boltzmann distributions as optimal models

To use relative entropy in a formulation of an EUR, we mainly need to specify the reference distributions  $\tilde{p}(x)$  and  $\tilde{q}(z)$ . Depending on the availability of side information about the actual distributions  $p(x)$  and  $q(z)$ , these distributions can be picked, such that they maximize an entropy. Let us make this idea more explicit by considering some examples in the following.

##### 1. No prior information

In the absence of any constraints, the optimal model, i.e., the model with maximum Shannon entropy, is given by a uniform distribution

$$\tilde{p}(x) = p_{\max}(x) = 1/d. \quad (15)$$

Then, the relative entropy between  $p(x)$  and  $p_{\max}(x)$  reduces to a simple difference of entropies

$$\begin{aligned} S(p\|p_{\max}) &= -S(p) + \ln d \\ &= -S(p) + S(p_{\max}), \end{aligned} \quad (16)$$

such that we can rewrite the Maassen-Uffink EUR (4) as

$$S(p\|p_{\max}) + S(q\|q_{\max}) \leq -\ln \frac{1}{c} - S(\rho) + 2 \ln d. \quad (17)$$

The sum of the divergencies of  $p(x)$  from  $p_{\max}(x)$  and  $q(x)$  from  $q_{\max}(x)$  is bounded from above. This bound becomes tighter the smaller the maximum overlap  $c$  in Eq. (5) is and the larger the von Neumann entropy  $S(\rho)$  is.

In the extreme case of a mutually unbiased basis (MUB) one has  $\ln(1/c) = \ln d$  and for a maximally mixed state  $S(\rho) = \ln d$ . In this case the right-hand side of Eq. (17) vanishes and neither  $p(x)$  nor  $q(x)$  can be distinguished from a uniform distribution with maximum entropy any more. This shows how the uncertainty relation has indeed an information theoretic significance in this formulation.

### 2. Given expectation values

Another convenient choice of constraints is a set of fixed expectation values. This is in particular interesting if we consider thermal states or in general if we have some additional macroscopic information about a set of observables, which are distributed according to  $p(x)$ . For example, we can consider the expectation value of the true distribution  $p(x)$

$$\mu_x = \sum_x p(x)x \quad (18)$$

and require that the reference distribution  $\tilde{p}(x)$  has the same mean value  $\tilde{\mu}_x = \mu_x$ . Then, the optimal model in the sense that  $S(\tilde{p})$  is maximal under the constraint  $\tilde{\mu}_x = \sum_x \tilde{p}(x)x = \mu_x$  is given by

$$\tilde{p}(x) = p_{\max}(x) = \frac{1}{Z_x} e^{-\gamma_x x}, \quad (19)$$

where  $\gamma_x$  is a Lagrange multiplier and  $Z_x$  is a constant for proper normalization. If there exists a set of observables  $Y_i$ , which commute with  $X$ , then we can include their expectation values, too. For example, if one of these operators is the Hamiltonian  $H$ , one can consider a thermal state as the optimal model. In particular, the condition of equal energy expectation values for the distributions  $p(x)$  and  $\tilde{p}(x) = p_{\max}(x)$  uniquely determines the inverse temperature  $\beta$  of the thermal model.

For a Boltzmann-type distribution  $p_{\max}(x)$  with equal expectation value we find for the relative entropy

$$\begin{aligned} S(p\|p_{\max}) &= -S(p) + \ln Z_x + \gamma_x \mu_x \\ &= -S(p) + S(p_{\max}), \end{aligned} \quad (20)$$

i.e., it again reduces to a difference of entropies. If we consider two observables  $X$  and  $Z$  and choose Boltzmann-type distributions as optimal models, a reformulation of the

Maassen-Uffink EUR (4) reads

$$\begin{aligned} S(p\|p_{\max}) + S(q\|q_{\max}) \\ \leq -\ln \frac{1}{c} - S(\rho) + \ln(Z_x Z_z) + \gamma_x \mu_x + \gamma_z \mu_z, \end{aligned} \quad (21)$$

where we used indices for all quantities related to the two different observables  $X$  and  $Z$ . These quantities are uniquely determined by the true distributions  $p(x)$  and  $q(z)$  and thus can be computed once the latter have been determined experimentally.

The latter considerations can be extended to the case where the mean values of the true distribution  $p(x)$  and the model  $\tilde{p}(x)$  are given, but do *not* agree  $\mu_x \neq \tilde{\mu}_x$ . In this case, the relative entropy acquires an additional term, which is proportional to the difference of the two,

$$S(p\|p_{\max}) = -S(p) + S(p_{\max}) + \gamma_x(\mu_x - \tilde{\mu}_x). \quad (22)$$

Nevertheless, the REUR (21) remains the same.

### 3. Given set of moments

We can generalize the latter considerations to the case where an arbitrary set of moments is constrained. More precisely, we may consider  $N$  functions  $m_1(x), \dots, m_N(x)$ , for which the expectation values

$$\langle m_j \rangle = \sum_x p(x) m_j(x) \quad (23)$$

are known. Thus the optimal model distribution  $\tilde{p}(x) = p_{\max}(x)$  needs to fulfill the set of constraints

$$\alpha_j(p, \tilde{p}) = \sum_x [\tilde{p}(x) - p(x)] m_j(x) = 0, \quad (24)$$

for  $j \in \{1, \dots, N\}$ . Assuming the existence of such an optimal model  $\tilde{p}(x)$ <sup>2</sup>, the maximum entropy principle dictates its form to be [49,51]

$$\tilde{p}(x) = p_{\max}(x) = \exp \left( \sum_{j=0}^N \lambda_j m_j(x) \right), \quad (25)$$

where  $\lambda_j$  for  $j > 0$  are Lagrangian multipliers ensuring the  $N$  constraints  $\alpha_j(p, \tilde{p}) = 0$  and  $\lambda_0$  is needed for proper normalization of the resulting optimal model distribution  $\tilde{p}(x)$ . Interestingly, the form of the solution (25) does not change if the constraints are given in the form of inequalities  $\alpha_j(p, \tilde{p}) \geq 0$ , but one has now additional constraints  $\lambda_j \geq 0$  for all  $j \in \{1, \dots, N\}$  for the optimization problem [52,53].

### C. Formulation solely in terms of relative entropies

Another interesting point is to rewrite (21) entirely in terms of relative entropies. To that end we need to replace the quantum entropy  $S(\rho)$  and the Shannon entropies of some optimal models by the quantum relative entropy, which is defined as [48,54]

$$S(\rho\|\tilde{\rho}) = -\text{Tr}\{\rho (\ln \rho - \ln \tilde{\rho})\}. \quad (26)$$

<sup>2</sup>Note that, in the case of continuous variables, solutions to this optimization problem often do not exist, especially when considering higher-order moments.



It can be considered as the quantum analog of (8) exhibiting the same properties on the level of density matrices  $\rho$  and  $\tilde{\rho}$ .

As for the classical relative entropies, we consider models of maximum entropy. In particular, we construct the measured state  $\rho_{X,\max}$  corresponding to the optimal classical distribution

$$\rho_{X,\max} = \sum_x p_{\max}(x) |x\rangle \langle x|, \quad (27)$$

which can be done analogously for  $\rho_{Z,\max}$ . Then, a quantum relative entropy of the actual state  $\rho$  with respect to such a model can be written as

$$S(\rho \| \rho_{X,\max}) = -S(\rho) + S(p_{\max}). \quad (28)$$

Furthermore, we can use that the uniform distribution

$$\rho_{\max} = \frac{1}{d} \mathbb{1} \quad (29)$$

allows one to rewrite the quantum entropy in a simple way

$$S(\rho) = -S(\rho \| \rho_{\max}) + \ln d. \quad (30)$$

Then, a formulation of (21) solely in terms of relative entropies reads

$$\begin{aligned} S(p \| p_{\max}) + S(q \| q_{\max}) \\ \leq \ln(cd) - S(\rho \| \rho_{\max}) + S(\rho \| \rho_{X,\max}) + S(\rho \| \rho_{Z,\max}). \end{aligned} \quad (31)$$

Note that for the maximum overlap as defined in Eq. (5) we have  $c \geq 1/d$  so that  $\ln(cd) \geq 0$ , with equality corresponding to mutually unbiased bases.

#### IV. CONTINUOUS VARIABLES

Now we turn to continuous variables, and specifically consider position  $X$  and momentum  $K$  in one spatial dimension. We assume the corresponding probability density functions  $f(x)$  and  $g(k)$  to be supported everywhere on the real line, if not stated differently.

Most importantly, the dimension of the Hilbert space is not finite any more and the optimal models of maximum entropy are described by probability density functions.

##### A. Position, momentum, and probability densities

In contrast to discrete variables, we cannot assign a meaningful quantum-classical state  $\rho_X$  after a measurement of  $X$  to our quantum system of interest. Mathematically, this is due to the fact that the eigenstates of  $X$  are not orthonormal (the same holds for  $K$ ). Instead, their overlaps are given by Dirac deltas

$$\langle x | x' \rangle = \delta(x - x'), \quad \langle k | k' \rangle = \delta(k - k'), \quad (32)$$

such that the formal extension of the projective measurement (12) leads to an operator

$$\rho \rightarrow \rho_X = \mathcal{X}(\rho) = \int dx f(x) |x\rangle \langle x|, \quad (33)$$

which is not trace class and thus cannot represent any quantum state. Nevertheless, one can still compute the density of the quantum state  $\rho$  in the eigenbasis of the observable  $X$ , such that the probability distribution density  $f(x) = \langle x | \rho | x \rangle$  [cf.

Eq. (10)] remains well defined. Then, the probability for obtaining an outcome between  $x$  and  $x + dx$  is given by  $f(x)dx$ .

For the following discussion it is important to look at the scalar product of position and momentum eigenstates

$$\langle x | k \rangle = \frac{1}{\sqrt{2\pi}} e^{ixk}, \quad (34)$$

which reflects the fact that the spectra of the two operators  $X$  and  $K$  are related by a Fourier transform. Since the latter is true for all positions  $x$  and momenta  $k$ , the maximum overlap  $c$  is given by

$$c = \max_{x,k} |\langle x | k \rangle|^2 = \frac{1}{2\pi} \quad (35)$$

and the two bases  $\{|x\rangle\}_x$  and  $\{|k\rangle\}_k$  can be considered as mutually unbiased bases in a continuous sense.

Let us note here that every experimental measurement of a position  $X$  or a momentum  $K$  can only be carried out with finite accuracies  $\delta x$  and  $\delta k$  leading formally to discrete probability distributions and measurement outcomes in the form of histograms.

More precisely, actual measurement outcomes can be considered as being sampled from the underlying probability density  $f(x)$ , such that one ends up with a discrete probability  $p_i$  to register for example a particle in the  $i$ th interval of size  $\delta x$  given by [6,9]

$$p_i = \int_{i\delta x}^{(i+1)\delta x} dx f(x). \quad (36)$$

Consequently, one can associate a Shannon entropy to this distribution,

$$S(p) = - \sum_i p_i \ln p_i, \quad (37)$$

which is a measure for the uncertainty about the measured distribution. In the limit of infinitely small bin sizes  $\delta x \rightarrow 0$ , the measured Shannon entropy  $S(p)$  diverges to  $+\infty$  as a consequence of allowing for an infinite and thus unphysical precision. Therefore, the true differential entropy of the underlying probability density  $S(f)$  can only be recovered in the limit  $\delta x \rightarrow 0$  after subtracting this infinite additive constant [6,9]

$$S(f) = \lim_{\delta x \rightarrow 0} [S(p) + \ln \delta x], \quad (38)$$

which shows again that the differential entropy  $S(f)$  is not the continuum limit of the Shannon entropy  $S(p)$  and that it does not inherit all of the Shannon entropies' properties. Nevertheless, it is possible to estimate the differential entropy  $S(f)$  from the measured probability distribution  $p_i$  [55].

Some of these problems might be avoided or relaxed by working with relative entropies instead of entropies. Moreover, it is of course interesting by itself to formulate EURs in terms of differential relative entropies, as we will do in the following.

##### B. Gaussian distributions as optimal models

In principle, one can consider the same constraints as before to obtain optimal models. For given expectation values

one would end up with distributions of Boltzmann type. Furthermore, if we put the quantum system into a box of finite length  $L$ , e.g.,  $x \in [-L/2, L/2]$ , we can choose the uniform distribution as optimal model, for the case that we do not have any constraints to implement. In the general case of an unbounded interval  $x \in (-\infty, \infty)$  this is not possible.<sup>3</sup>

Of special interest for the continuous case is a Gaussian distribution

$$\tilde{f}(x) = f_{\max}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right), \quad (39)$$

which represents the optimal model, in the sense of a maximum differential entropy  $S(f)$ , for known variance

$$\sigma_x^2 = \int dx f(x)(x - \mu_x)^2 \quad (40)$$

and mean

$$\mu_x = \int dx f(x)x. \quad (41)$$

Computing the entropy of the Gaussian distribution (39) gives

$$S(f_{\max}) = \frac{1}{2} \ln(2\pi e\sigma_x^2), \quad (42)$$

such that the reformulation of the Frank-Lieb EUR (7) in terms of relative entropies reads

$$S(f\|f_{\max}) + S(g\|g_{\max}) \leq -S(\rho) + 1 + \ln(\sigma_x\sigma_k). \quad (43)$$

Again the right-hand side is fully determined by the true distributions  $f(x)$  and  $g(k)$ .

Similarly, we can reformulate the EUR by Białynicki-Birula and Mycielski (2),

$$S(f\|f_{\max}) + S(g\|g_{\max}) \leq \ln 2 + \ln(\sigma_x\sigma_k). \quad (44)$$

This relation is particularly interesting if we express it in terms of variances again

$$\sigma_x\sigma_k \geq \frac{1}{2} e^{S(f\|f_{\max})+S(g\|g_{\max})}. \quad (45)$$

Compared to Robertson's formulation (1), the latter uncertainty relation provides a stronger bound for non-Gaussian distributions, where non-Gaussianity is measured in terms of relative entropies. Hence it becomes visible in which cases the EUR (2) is stronger than the uncertainty relation (1). The latter relation was already discussed in [16] (see also [10]), where the relative entropy with respect to optimal Gaussians  $S(f\|f_{\max})$  was interpreted as the neg-entropy  $\mathcal{J}(f) = S(f_{\max}) - S(f)$  accounting for non-Gaussianity.<sup>4</sup>

Finally, let us briefly comment on a possible rewriting of (43) entirely in terms of relative entropies as done in (31) for discrete variables. The strategy used in the discrete case breaks down for continuous variables due to the nonexistence of a measured state in Eq. (33). We did not find a way to circumvent this problem but will argue in Sec. V that one arrives at a satisfactory formulation nevertheless.

<sup>3</sup>Note that in the case of  $x \in [-L/2, L/2]$  we still can have (discrete)  $k \in (-\infty, \infty)$ , such that the uniform model may not be chosen for both distributions at once.

<sup>4</sup>See also [56] for a discussion concerning non-Gaussianity of quantum states measured in terms of *quantum* relative entropy.

## V. UNCERTAINTY RELATION BASED ON RELATIVE ENTROPY

In the last section we generalize the two cases of discrete and continuous variables and provide a REUR which covers both cases. Therefore, in the following discussion we will write  $p(x)$  and  $q(z)$  for *every* type of distribution and their precise meanings have to be deduced from the context. Furthermore, we discuss the resulting REUR in more detail.

### A. Model distributions of maximum entropy

In the two previous sections we were mainly concerned with finding suitable model distributions  $\tilde{p}(x)$  and  $\tilde{q}(x)$ , which turned out to be the key step for reformulating statements with entropy in terms of relative entropy (cf. also Refs. [57] and [58]).

For our purposes the most useful models  $\tilde{p}(x)$  and  $\tilde{q}(x)$  were the ones which allowed one to write a relative entropy as a simple difference of entropies. Since in general relative entropy can be written as

$$S(p\|\tilde{p}) = -S(p) + S(p, \tilde{p}), \quad (46)$$

i.e., as the difference between a cross entropy<sup>5</sup>

$$S(p, \tilde{p}) = - \sum_x p(x) \ln \tilde{p}(x) \quad (47)$$

and an entropy  $S(p)$ , we are interested in finding models  $\tilde{p}(x)$  for which the cross entropy  $S(p, \tilde{p})$  reduces to the entropy of the model  $S(\tilde{p})$  under certain additional constraints, i.e.,

$$S(p, \tilde{p}) = S(\tilde{p}) \quad (48)$$

for a set of constraints

$$\alpha_j(p, \tilde{p}) = 0, \quad (49)$$

where  $j \in \{1, \dots, N\}$  and  $N$  is typically a small natural number. Examples for convenient constraints included equal standard deviation  $\sigma_x = \tilde{\sigma}_x$  or equal expectation values of macroscopic quantities  $\langle O \rangle_p = \langle O \rangle_{\tilde{p}}$ .

The condition (48) can be inserted into the relative entropy, which simply gives

$$S(p\|\tilde{p}) = -S(p) + S(\tilde{p}). \quad (50)$$

Then, the non-negativity of the relative entropy is equivalent to

$$S(p) \leq S(\tilde{p}). \quad (51)$$

Since the latter inequality has to hold for all true distributions  $p(x)$ , the condition (48) requires that the model is *optimal*  $\tilde{p}(x) = p_{\max}(x)$ , in the sense that it maximizes the classical entropy under the given constraints.

Additionally, we can consider the case where the condition (48) is released below in the sense that we allow for terms linear in the constraints. This still leads to the distribution of maximum entropy, but it is less optimal in the sense that some of its properties do not coincide with those of the true

<sup>5</sup>For continuous distributions the sum has to be replaced by an integral  $\int dx$ .

distributions, for example, they could have different mean values  $\mu_x \neq \tilde{\mu}_x$ .

Another advantage of optimal model distributions is that they often fulfill the support condition  $\text{supp}[p(x)] \subseteq \text{supp}[p_{\max}(x)]$  for all  $x$ , such that the relative entropy  $S(p\|p_{\max})$  will be finite. As a consequence, a sum of relative entropies is bounded from above in a trivial way,

$$S(p\|p_{\max}) + S(q\|q_{\max}) \leq S(p_{\max}) + S(q_{\max}), \quad (52)$$

which can be seen from combining (46) with (48) and using that

$$S(p) + S(q) \geq 0 \quad (53)$$

also holds for the continuous case. Formulating an entropic uncertainty relation in terms of relative entropies is then equivalent to finding a smaller bound than (52).

### B. Uncertainty relation and discussion

Denoting the model distributions with maximum entropy and equal constraints by  $p_{\max}(x)$  and  $q_{\max}(z)$  allows us to state the *relative entropic uncertainty relation* (REUR) for the two true distributions  $p(x)$  and  $q(z)$  in its most general form

$$S(p\|p_{\max}) + S(q\|q_{\max}) \leq -\ln \frac{1}{c} - S(\rho) + S(p_{\max}) + S(q_{\max}), \quad (54)$$

where  $c$  is again the maximum overlap between any two eigenstates in (5) for projective measurements, or the generalization in Eq. (6) for POVMs.

One has to be careful with the interpretation of this inequality because the usual logic is in some sense reversed. Therefore, a few comments are in order.

(a) *Sum of relative entropies is bounded from above.* The nontrivial bound on a sum of relative entropies comes from above and not from below. In addition, a sum of relative entropies is bounded from below by zero due to the non-negativity of all summands. Furthermore, it is bounded from above as shown in Eq. (52). The upper bound tells that the ability to distinguish the measurement outcome distributions  $p(x)$  and  $q(x)$  from models with minimal information is bounded. The discriminating power on the right-hand side of (54) gets smaller when the overlap between basis functions in Eq. (5) gets smaller, or when the von Neumann entropy gets larger.

Being able to distinguish a distribution well from a model means to have a high amount of information additional to the prior information encoded in the model. Since all uncertainty relations are statements about never having all information about two noncommuting observables, this translates to an upper bound for a sum of relative entropies with respect to models representing maximal missing information.

(b) *Smaller bound from smaller overlap and mixedness.* The bound is smaller if the two observables have a smaller quantum incompatibility measured in terms of  $c$ . It is also smaller if the state of the system becomes more mixed due to the term  $-S(\rho) \leq 0$ . Thus *quantum incompatibility* or *mixedness* is encoded in having a *small bound*. Note that the right-hand side of (54) becomes largest, and the bound correspondingly weakest, when  $c \rightarrow 1$  and  $S(\rho) \rightarrow 0$ . This

corresponds to the classical situation and absence of statistical fluctuations in the state  $\rho$ . Nevertheless, the discriminating power on the left-hand side of (54) is then still bounded, namely by the sum of maximum entropies [cf. Eq. (52)].

(c) *Maassen-Uffink and Frank-Lieb relations follow as special cases.* As we have seen in Secs. III and IV, the REUR (54) reduces to the Maassen-Uffink relation (4) for discrete and to the Frank-Lieb relation (7) for continuous variables, respectively. Therefore, the tightness of the REUR (54) depends on the type of variables under consideration. For discrete variables, it is tight for mutually unbiased bases and states that are diagonal in either one of the two bases. In contrast, for continuous variables it becomes tight in the infinite temperature limit. Furthermore, as the REUR (54) is mathematically equivalent to the measure-theoretic formulation by Frank and Lieb put forward in [21], it could be proven in complete analogy using the Golden-Thompson inequality and Gibbs variational principle.

(d) *Discrete and continuous spectra are unified.* The main advantage of the REUR (54) is that the left-hand side is well defined in both cases, i.e., for operators with discrete as well as continuous spectrum, which is one of the nice features of relative entropy. On the right-hand side we still have a sum of entropies, but for optimal models  $p_{\max}(x)$  and  $q_{\max}(z)$  we can always express these in terms of quantities appearing in the constraints  $\alpha_j$ . Thus the relation is indeed free of ambiguities.

(e) *Bound contains additional knowledge.* In most cases, the bound of an uncertainty relation depends on some measure of quantum incompatibility (either a commutator or an overlap) and the state  $\rho$ . Interestingly, the bound in the REUR (54) does also depend on the additional knowledge encoded in the optimal models  $p_{\max}(x)$  and  $q_{\max}(z)$ , i.e., on the constraints. Therefore, any available information about the distributions can be implemented directly in the presented uncertainty relation. This feature may be of particular interest for experimental applications where different kinds of constraints may be accessed.

(f) *Change of normalization.* Specifically for continuous variables it is instructive to study scaling transformations of the form  $|x\rangle \rightarrow |x'\rangle = \alpha|x\rangle$ . For  $|\alpha|^2 \neq 1$  they change the normalization of the basis  $\{|x\rangle\}_x$ . One needs to complement this change of basis with a change of integration measure  $dx \rightarrow dx' = |\alpha|^{-2}dx$  such that the probability

$$\langle x|\rho|x\rangle dx \rightarrow \langle x'|\rho|x'\rangle dx' = \langle x|\rho|x\rangle dx \quad (55)$$

remains unchanged. The differential entropy changes then according to  $S(p_{\max}) \rightarrow S(p_{\max}) - \ln(|\alpha|^2)$ . For the maximum overlap as defined in (5) one has  $c \rightarrow c' = |\alpha|^2 c$  such that  $-\ln(1/c) \rightarrow -\ln(1/c') = -\ln(1/c) + \ln(|\alpha|^2)$ . Taken together, the right-hand side of (54) is invariant under such a “change of normalization” transformation. When  $|z\rangle$  corresponds to a continuous variable there is a similar invariance.

### C. Example of angle and angular momentum

We close our analysis by considering a nontrivial example, namely angular momentum states and corresponding discrete or continuous angles. We will demonstrate here how the bound of the REUR (54) behaves if we start with operators with discrete spectra and take the continuum and infinite vol-

ume limits. Note that the limits have to be understood in a formal way and that the left-hand side of (54) is always well defined as explained around (9).

We start from a finite set of  $2J + 1$  spin or angular momentum eigenstates  $|m\rangle$  with  $L_z|m\rangle = m|m\rangle$  and where  $m \in \{-J, \dots, J\}$  is an integer. As usual,  $J$  can be integer or half-integer. Following Ref. [59] we introduce now the angle states

$$|\phi\rangle = \frac{1}{\sqrt{2J+1}} \sum_{m=-J}^J e^{-im\phi} |m\rangle. \quad (56)$$

In particular, one has with this definition  $|\phi\rangle = |\phi + 2\pi\rangle$  and

$$e^{-i\phi L_z} |\phi\rangle = |\phi + \varphi\rangle, \quad (57)$$

as it should be. Because there is a continuum of angle states, they are overcomplete and have the overlap

$$\langle\varphi|\phi\rangle = \frac{\sin[(J + \frac{1}{2})(\varphi - \phi)]}{(2J+1) \sin[\frac{1}{2}(\varphi - \phi)]}. \quad (58)$$

In this regard the angle states are similar to coherent states. Note also the completeness relation

$$\frac{2J+1}{2\pi} \int_0^{2\pi} d\phi |\phi\rangle \langle\phi| = \sum_{m=-J}^J |m\rangle \langle m| = \mathbb{1}. \quad (59)$$

One can therefore understand a measurement of the continuous angle  $\phi$  as a POVM.

It is also possible to restrict to a discrete set of angle states  $|\theta_j\rangle$  with

$$\theta_j = \theta_0 + \frac{2\pi j}{2J+1}, \quad j = 0, 1, \dots, 2J. \quad (60)$$

These discrete angle states are actually orthonormal,  $\langle\theta_j|\theta_k\rangle = \delta_{jk}$ . In the continuum limit  $J \rightarrow \infty$  the corresponding angles become dense but the states need then to be normalized as continuum states. It is also possible to introduce position variables on the circle  $x = 2\pi R\phi$ , conjugate momenta  $k = L_z/(2\pi R)$ , and to consider an “infinite volume” limit  $R \rightarrow \infty$ .

Now let us discuss our relative entropic uncertainty relations in this context. First, for a measurement of the angular momentum  $L_z$  with outcome  $m$  one can use in the case of finite  $J$  any of the reference distributions discussed in Sec. III. For example, if no information is available the uniform distribution would be a sensible reference, and if  $\langle m \rangle$  and  $\langle m^2 \rangle$  are known, a sensible reference distribution would be of the form

$$q_{\max}(m) = \exp(\lambda_0 + \lambda_1 m + \lambda_2 m^2), \quad (61)$$

with Lagrangian multipliers  $\lambda_j$ . The corresponding entropy appearing on the right-hand side of Eq. (54) is simply  $S(q_{\max}) = -\lambda_0 - \lambda_1 \langle m \rangle - \lambda_2 \langle m^2 \rangle$ . The Gaussian reference distribution (61) remains (for  $\lambda_2 < 0$ ) also normalizable in the continuum limit  $J \rightarrow \infty$  and it corresponds to a continuous Gaussian distribution of momenta  $k = m/(2\pi R)$  in the infinite volume limit  $R \rightarrow \infty$  when the corresponding variances are held fixed.

For a measurement of the angle the situation is more involved. First, for finite  $J$  one can consider the projective measurement in the basis  $|\theta_j\rangle$  corresponding to the angles (60). This is then a discrete observable which can be treated

similar to  $L_z$  above. Note that the overlap as defined in Eq. (5) is minimal,  $c = \max_{m,j} |\langle m|\theta_j\rangle|^2 = 1/(2J+1)$ .

A measurement of the continuous angle states  $|\phi\rangle$  is instead a POVM. Because the range of possible values is finite, the uniform distribution  $p_{\max}(\phi) = 1/(2\pi)$  with  $S(p_{\max}) = \ln(2\pi)$  is an admissible reference distribution with maximum entropy when no additional information is given. Another interesting reference distribution is the von Mises distribution [60]

$$p_{\max}(\phi) = \frac{e^{\kappa \cos(\phi - \mu)}}{2\pi I_0(\kappa)}, \quad (62)$$

where  $I_i(\kappa)$  is the modified Bessel function of the first kind and of order  $i$ . It corresponds to a maximum entropy distribution for  $\phi$  under the condition that the first circular moment

$$\langle e^{i\phi} \rangle = \frac{I_1(\kappa)}{I_0(\kappa)} e^{i\mu} \quad (63)$$

is known. The magnitude and complex phase of the latter fix the parameters  $\kappa \geq 0$  and  $0 \leq \mu < 2\pi$ . The differential entropy is in this case given by

$$S(p_{\max}) = \ln[2\pi I_0(\kappa)] - \kappa \frac{I_1(\kappa)}{I_0(\kappa)}, \quad (64)$$

and the quantum incompatibility  $c$  as defined in Eq. (6) evaluates to  $c = 1/(2\pi)$ . We have dropped here a multiplicative term  $d\phi$  which would appear from a direct application of Eq. (59) in Eq. (6) but cancels with a similar term that arises in the transition from a Shannon entropy to the differential entropy in Eq. (64). This fixes all state independent terms on the right-hand side of the REUR (54).

Let us note that, for large  $\kappa$ , the von Mises distribution becomes strongly localized around  $\phi = \mu$  and approaches a Gaussian shape. In this sense one can recover the Gaussian models discussed in Sec. IV B in the infinite volume limit  $R \rightarrow \infty$  for fixed variance of position  $x$ .

We also note that, up to the different normalization of states, the two possibilities to look at angle measurements (discrete as PVM and continuous as POVM) become equivalent in the continuum limit  $J \rightarrow \infty$ .

## VI. CONCLUSION AND OUTLOOK

In summary, we have investigated a formulation of an entropic uncertainty relation in terms of *relative entropies*. In particular, we found that a sum of relative entropies of the two true distributions with respect to model distributions of maximum entropy are bounded from above in a nontrivial way. More precisely, there is a bound depending on the entropies of the optimal models, which gets reduced by *quantum incompatibility* in terms of the maximum overlap  $c$  and *mixedness* in terms of the von Neumann entropy  $S(\rho)$  of the quantum state  $\rho$ . The main advantage of the presented formulation was that it allowed one to cover the cases of observables with discrete and continuous spectra within one and the same entropic uncertainty relation. This was due to the fact that relative entropy, in contrast to Shannon entropy, behaves well when considering the continuum limit.



We have illustrated the benefits of this formulation on the specific example of finite and infinite spaces of angular momentum states and the conjugate angle variables. All state-independent quantities appearing in the uncertainty relation can then be easily evaluated and behave also favorable under the continuum as well as infinite volume limits.

For the future it might be interesting to generalize the presented approach to situations with quantum memory.

Another interesting point for future work concerns the application of the (relative) entropic uncertainty relation to quantum systems where entropies exhibit (unphysical) infinities while relative entropies remain finite. What we have in

mind here are specifically applications to quantum field theory in terms of (functional) relative entropies. Such a formulation would allow one to quantify entropic uncertainty, e.g., for a scalar quantum field and its conjugate momentum field.

### ACKNOWLEDGMENTS

This work is supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy EXC 2181/1-390900948 (the Heidelberg STRUCTURES Excellence Cluster) and under 273811115 SFB 1225 ISOQUANT, as well as FL 736/3-1.

- 
- [1] W. Heisenberg, Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik, *Z. Phys.* **43**, 172 (1927).
  - [2] E. H. Kennard, Zur Quantenmechanik einfacher Bewegungstypen, *Z. Phys.* **44**, 326 (1927).
  - [3] H. Weyl, *Gruppentheorie und Quantenmechanik* (Hirzel, Leipzig, 1928).
  - [4] H. P. Robertson, The uncertainty principle, *Phys. Rev.* **34**, 163 (1929).
  - [5] E. Schrödinger, Zum Heisenbergschen Unschärfepinzipp, Sitz. Preuß. Akad. Wiss. Physik.-math. Kl. **14**, 296 (1930).
  - [6] I. Białynicki-Birula and Ł. Rudnicki, Entropic uncertainty relations in quantum physics, *Statistical Complexity* (Springer, Dordrecht, 2011).
  - [7] C. E. Shannon, A mathematical theory of communication, *Bell Syst. Tech. J.* **27**, 379 (1948).
  - [8] I. Białynicki-Birula, Entropic uncertainty relations in quantum mechanics, in *Quantum Probability and Applications II*, edited by L. Accardi and W. von Waldenfels (Springer, New York, 1984), pp. 90–103.
  - [9] P. J. Coles, M. Berta, M. Tomamichel, and S. Wehner, Entropic uncertainty relations and their applications, *Rev. Mod. Phys.* **89**, 015002 (2017).
  - [10] A. Hertz and N. J. Cerf, Continuous-variable entropic uncertainty relations, *J. Phys. A: Math. Theor.* **52**, 173001 (2019).
  - [11] S. Floerchinger and J.-M. Schwindt, Neutrino flavor-mass uncertainty relations and an entanglement-assisted determination of the PMNS matrix, *Phys. Rev. D* **102**, 093001 (2020).
  - [12] H. Everett, “Relative state” formulation of quantum mechanics, *Rev. Mod. Phys.* **29**, 454 (1957).
  - [13] I. I. Hirschman, A note on entropy, *Am. J. Math.* **79**, 152 (1957).
  - [14] W. Beckner, Inequalities in Fourier analysis, *Ann. Math.* **102**, 159 (1975).
  - [15] I. Białynicki-Birula and J. Mycielski, Uncertainty relations for information entropy in wave mechanics, *Commun. Math. Phys.* **44**, 129 (1975).
  - [16] W. Son, Role of quantum non-Gaussian distance in entropic uncertainty relations, *Phys. Rev. A* **92**, 012114 (2015).
  - [17] D. Deutsch, Uncertainty in Quantum Measurements, *Phys. Rev. Lett.* **50**, 631 (1983).
  - [18] H. Maassen and J. B. M. Uffink, Generalized Entropic Uncertainty Relations, *Phys. Rev. Lett.* **60**, 1103 (1988).
  - [19] K. Kraus, Complementary observables and uncertainty relations, *Phys. Rev. D* **35**, 3070 (1987).
  - [20] 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 (2010).
  - [21] R. L. Frank and E. H. Lieb, Entropy and the uncertainty principle, *Ann. Henri Poincaré* **13**, 1711 (2012).
  - [22] M. Krishna and K. R. Parthasarathy, An entropic uncertainty principle for quantum measurements, *Ind. J. Stat.* **64**, 842 (2002).
  - [23] M. Tomamichel, A framework for non-asymptotic quantum information theory, Ph.D. thesis, Department of Physics, ETH Zurich, 2012, <https://doi.org/10.3929/ethz-a-7356080>.
  - [24] P. J. Coles and M. Piani, Improved entropic uncertainty relations and information exclusion relations, *Phys. Rev. A* **89**, 022112 (2014).
  - [25] M. M. Wilde, *Quantum Information Theory* (Cambridge University Press, Cambridge, UK, 2013).
  - [26] M. Grabowski, Wehrl-Lieb's inequality for entropy and the uncertainty relation, *Rep. Math. Phys.* **20**, 153 (1984).
  - [27] M. J. W. Hall, Almost-periodic time observables for bound quantum systems, *J. Phys. A: Math. Theor.* **41**, 255301 (2008).
  - [28] A. Boette, R. Rossignoli, N. Gigena, and M. Cerezo, System-time entanglement in a discrete-time model, *Phys. Rev. A* **93**, 062127 (2016).
  - [29] M. J. W. Hall, Entropic Heisenberg limits and uncertainty relations from the Holevo information bound, *J. Phys. A: Math. Theor.* **51**, 364001 (2018).
  - [30] P. J. Coles, V. Katariya, S. Lloyd, I. Marvian, and M. M. Wilde, Entropic Energy-Time Uncertainty Relation, *Phys. Rev. Lett.* **122**, 100401 (2019).
  - [31] I. Białynicki-Birula, Formulation of the uncertainty relations in terms of the Rényi entropies, *Phys. Rev. A* **74**, 052101 (2006).
  - [32] J. M. Renes and J.-C. Boileau, Conjectured Strong Complementary Information Tradeoff, *Phys. Rev. Lett.* **103**, 020402 (2009).
  - [33] M. Tomamichel and R. Renner, Uncertainty Relation for Smooth Entropies, *Phys. Rev. Lett.* **106**, 110506 (2011).
  - [34] P. J. Coles, R. Colbeck, L. Yu, and M. Zvolak, Uncertainty Relations from Simple Entropic Properties, *Phys. Rev. Lett.* **108**, 210405 (2012).
  - [35] R. L. Frank and E. H. Lieb, Extended quantum conditional entropy and quantum uncertainty inequalities, *Commun. Math. Phys.* **323**, 487 (2013).
  - [36] M. Rumin, Balanced distribution-energy inequalities and related entropy bounds, *Duke Math. J.* **160**, 567 (2011).

- [37] M. Rumin, An entropic uncertainty principle for positive operator valued measures, *Lett. Math. Phys.* **100**, 291 (2012).
- [38] S. Kullback and R. A. Leibler, On information and sufficiency, *Ann. Math. Stat.* **22**, 79 (1951).
- [39] S. Kullback, *Information Theory and Statistics* (Dover Publications, Mineola, NY, 1968).
- [40] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, second ed. (John Wiley and Sons, New York, 2006).
- [41] 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).
- [42] M. J. W. Hall, Information Exclusion Principle for Complementary Observables, *Phys. Rev. Lett.* **74**, 3307 (1995).
- [43] M. J. W. Hall, Quantum information and correlation bounds, *Phys. Rev. A* **55**, 100 (1997).
- [44] A. Barchielli, M. Gregoratti, and A. Toigo, Measurement uncertainty relations for position and momentum: Relative entropy formulation, *Entropy* **19**, 301 (2017).
- [45] A. Barchielli, M. Gregoratti, and A. Toigo, Measurement uncertainty relations for discrete observables: Relative entropy formulation, *Commun. Math. Phys.* **357**, 1253 (2018).
- [46] P. Busch, P. Lahti, and R. F. Werner, Colloquium: Quantum root-mean-square error and measurement uncertainty relations, *Rev. Mod. Phys.* **86**, 1261 (2014).
- [47] 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).
- [48] V. Vedral, The role of relative entropy in quantum information theory, *Rev. Mod. Phys.* **74**, 197 (2002).
- [49] E. T. Jaynes, Information theory and statistical mechanics (notes by the lecturer), *Statistical Physics 3* (W. A. Benjamin, New York, 1963), pp. 181–218.
- [50] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary Edition* (Cambridge University Press, Cambridge, UK, 2010).
- [51] E. T. Jaynes, Information theory and statistical mechanics, *Phys. Rev.* **106**, 620 (1957).
- [52] 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).
- [53] P.-T. de Boer, D. P. Kroese, S. Mannor, and R. Y. Rubinstein, A tutorial on the cross-entropy method, *Ann. Op. Res.* **134**, 19 (2005).
- [54] H. Umegaki, Conditional expectation in an operator algebra. IV. Entropy and information, *Kodai Math. Seminar Rep.* **14**, 59 (1962).
- [55] J. Beirlant, E. J. Dudewicz, L. Györfi, and E. C. van der Meulen, Nonparametric entropy estimation: An overview, *Int. J. Math. Stat. Sci.* **6**, 17 (1997).
- [56] 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(R) (2008).
- [57] S. Floerchinger and T. Haas, Thermodynamics from relative entropy, *Phys. Rev. E* **102**, 052117 (2020).
- [58] N. Dowling, S. Floerchinger, and T. Haas, Second law of thermodynamics for relativistic fluids formulated with relative entropy, *Phys. Rev. D* **102**, 105002 (2020).
- [59] S. M. Barnett and D. T. Pegg, Quantum theory of rotation angles, *Phys. Rev. A* **41**, 3427 (1990).
- [60] K. V. Mardia and P. E. Jupp, *Directional Statistics* (Wiley, New York, 2009).