Entropic Uncertainty Relations and their Applications

Patrick J. Coles*

Institute for Quantum Computing and Department of Physics and Astronomy, University of Waterloo, N2L3G1 Waterloo, Ontario, Canada

Mario Berta†

Institute for Quantum Information and Matter, California Institute of Technology, Pasadena, CA 91125, USA

Marco Tomamichel‡

School of Physics, The University of Sydney, Sydney, NSW 2006, Australia

Stephanie Wehner§

QuTech, Delft University of Technology, 2628 CJ Delft, Netherlands

Heisenberg's uncertainty principle forms a fundamental element of quantum mechanics. Uncertainty relations in terms of entropies were initially proposed to deal with conceptual shortcomings in the original formulation of the uncertainty principle, and hence play an important role in quantum foundations. More recently, entropic uncertainty relations have emerged as the central ingredient in the security analysis of almost all quantum cryptographic protocols, ranging from quantum key distribution to two-party quantum cryptography. This review surveys entropic uncertainty relations that capture Heisenberg's idea that the results of incompatible measurements are impossible to predict, covering both finite- and infinite-dimensional measurements. These ideas are then extended to incorporate quantum correlations between the observed object and its environment, allowing for a variety of recent, more general formulations of the uncertainty principle. Finally various applications are discussed, ranging from entanglement witnessing to wave-particle duality to quantum cryptography.

CONTENTS

* pcoles@uwaterloo.ca
† berta@caltech.edu
‡ marco.tomamichel@sydney.edu.au
§ s.d.c.wehner@tudelft.nl

I. INTRODUCTION

Quantum mechanics has revolutionized our understanding of the world around us. Relative to classical mechanics, the most revolutionary change in our understanding is that the quantum world — our world — is inherently unpredictable.

By far the most famous statement of unpredictability is Heisenberg's uncertainty principle (Heisenberg, 1927), which we treat here as a statement about preparation uncertainty. Roughly speaking, it states that it is impossible to prepare a quantum particle for which both position and momentum are sharply defined. Operationally, consider a source that consistently prepares copies of a quantum particle in the same way, as shown in Fig. 1. For each copy, suppose we randomly measure either its position or its momentum (but we never attempt to measure

both quantities for the same particle[1]). We record the outcomes and sort them into two sequences associated with the two different measurements. The uncertainty principle states that it is impossible to predict both the outcome of the position and the momentum measurements: at least one of the two sequences of outcomes will be unpredictable. More precisely, the better such a preparation procedure allows one to predict the outcome of the position measurement, the more uncertain the outcome of the momentum measurement will be, and vice versa.

A beautiful aspect of quantum mechanics is that it allows for simple quantitative statements of this idea, i.e., constraints on the predictability of observable pairs like position and momentum. These quantitative statements are known as uncertainty relations. It is worth noting that Heisenberg's original argument, while conceptually enlightening, was heuristic. The first, rigorously-proven uncertainty relation for position Q and momentum P is due to Kennard (1927). It establishes that (see also the work of Weyl (1928))

$$\sigma_Q \sigma_P \geq \frac{\hbar}{2}\,, \tag{1}$$

where σ_Q and σ_P denote the standard deviation of the position and momentum, respectively, and $\hbar$ is the reduced Planck constant.

We now know that Heisenberg's principle applies much more generally, not only to position and momentum. Other examples of pairs of observables observing an uncertainty relation include the phase and excitation number of a harmonic oscillator, the angle and the orbital angular momentum of a particle, and orthogonal components of spin angular momentum. In fact, for arbitrary observables[2], X and Z, Robertson (1929) showed that

$$\sigma_X \sigma_Z \geq \frac{1}{2}|\langle\psi|[X,Z]|\psi\rangle|\,, \tag{2}$$

where $[\cdot,\cdot]$ denotes the commutator. Note a distinct difference between (1) and (2): the right-hand side of the former is a constant whereas that of the latter can be state-dependent, an issue that we will discuss more in Sec. II.

These relations have a beauty to them and also give conceptual insight. Equation (1) identifies $\hbar$ as a fundamental limit to our knowledge. More generally (2) identifies the commutator as the relevant quantity for determining how large the knowledge tradeoff is for two observables. One could argue that a reasonable goal in our studies of uncertainty in quantum mechanics should be to find simple, conceptually insightful statements like this, if they exist.

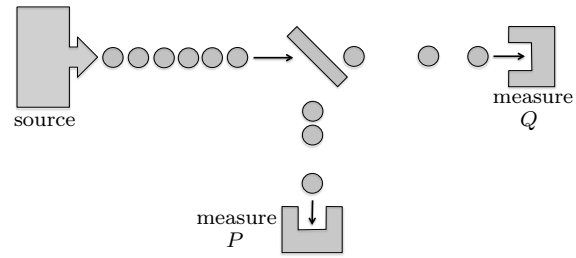


FIG. 1 Physical scenario relevant to preparation uncertainty relations. Each incoming particle is randomly measured using either measurement P or measurement Q. An uncertainty relation says we cannot predict the outcomes of both P and Q. If we can predict the outcome of P well, then we are necessarily uncertain if measurement Q would have been performed and vice versa.

If this problem was only of fundamental importance, it would be a well-motivated one. Yet in recent years there is new motivation to study the uncertainty principle. The rise of quantum information theory has led to new applications of quantum uncertainty, for example in quantum cryptography. In particular quantum key distribution is already commercially marketed and its security crucially relies on Heisenberg's uncertainty principle. (We will discuss various applications in Sec. VI.) There is a clear need for uncertainty relations that are directly applicable to these technologies.

In the above uncertainty relations, (1) and (2), uncertainty has been quantified using the standard deviation of the measurement results. This is, however, not the only way to express the uncertainty principle. It is instructive to consider what preparation uncertainty means in the most general setting. Suppose we have prepared a state ρ on which we can perform two (or more) possible measurements labeled by θ. Let us use x to label the outcomes of such measurement. We can then identify a list of (conditional) probabilities

$$S_\rho = \left\{p(x|\theta)_\rho\right\}_{x,\theta}\,, \tag{3}$$

where $p(x|\theta)_\rho$ denotes the probability of obtaining measurement outcome x when performing the measurement θ on the state ρ. Quantum mechanics predicts restrictions on the set S_ρ of allowed conditional probability distributions that are valid for all or a large class of states ρ. Needless to say, there are many ways to formulate such restrictions on the set of allowed distributions.

In particular, information theory offers a very versatile abstract framework that allows us to formalize notions like uncertainty and unpredictability. This theory is the basis of modern communication technologies and cryptography and has been successfully generalized to include

[1] Section I.A briefly notes other uncertainty principles that involve consecutive or joint measurements.

[2] More precisely, Robertson's relation refers to observables with bounded spectrum.

quantum effects. The preferred mathematical quantity to express uncertainty in information theory is entropy. Entropies are functionals on random variables and quantum states that aim to quantify their inherent uncertainty. Amongst a myriad of such measures, we mainly restrict our attention to the Boltzmann–Gibbs–Shannon entropy (Boltzmann, 1872; Gibbs, 1876; Shannon, 1948) and its quantum generalization, the von Neumann entropy (von Neumann, 1932). Due to their importance in quantum cryptography, we will also consider Rényi entropic measures (Rényi, 1961) such as the min-entropy. Entropy is a natural measure of uncertainty, perhaps even more natural than the standard deviation, as we argue in Sec. II.

Can the uncertainty principle be formulated in terms of entropy? This question was first brought up by Everett (1957) and answered in the affirmative by Hirschman (1957) who considered the position and momentum observables, formulating the first *entropic uncertainty relation*. This was later improved by Beckner (1975) and Białynicki-Birula and Mycielski (1975), who obtained the relation[3]

$$h(Q) + h(P) \geq \log(e\pi\hbar)\,, \tag{4}$$

where h is the differential entropy (defined in (7) below). Białynicki-Birula and Mycielski (1975) also showed that (4) is stronger than, and hence implies, Kennard's relation (1).

The extension of the entropic uncertainty relation to observables with finite spectrum[4] was given by Deutsch (1983), and later improved by Maassen and Uffink (1988) following a conjecture by Kraus (1987). The result of Maassen and Uffink (1988) is arguably the most well-known entropic uncertainty relation. It states that

$$H(X) + H(Z) \geq \log \frac{1}{c}, \tag{5}$$

where H is Shannon's entropy, and c denotes the maximum overlap between any two eigenvectors of the X and Z observables. Just as (2) established the commutator as an important parameter in determining the uncertainty tradeoff for standard deviation, (5) established the maximum overlap c as a central parameter in entropic uncertainty.

While these articles represent the early history of entropic uncertainty relations, there has recently been an explosion of work on this topic. One of the most important recent advances concerns a generalization of the uncertainty paradigm that allows the measured system to be correlated to its environment in a non-classical way. Entanglement between the measured system and the environment can be exploited to reduce the uncertainty of an observer (with access to the environment) below the usual bounds.

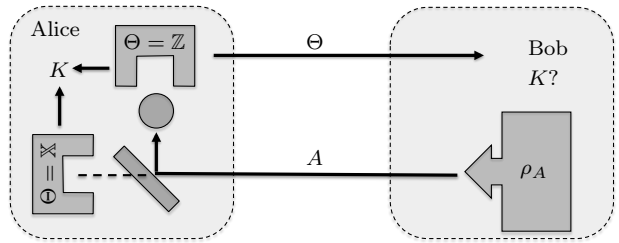


FIG. 2 Cartoon showing a guessing game with players Alice and Bob. First, Bob prepares A in state ρ_A and sends it to Alice. Second Alice measures either $\mathbb{X}$ or $\mathbb{Z}$ with equal probability and stores the measurement choice in bit Θ. Third, Alice stores the measurement outcome in bit K and reveals the measurement choice Θ to Bob. Bob wins the game if he can guess K (given Θ). Entropic uncertainty relations like the Maassen-Uffink relation (5) can be understood as fundamental constraints on the optimal guessing probability.

To explain this extension, let us introduce a modern formulation of the uncertainty principle as a so-called guessing game, which makes such extensions of the uncertainty principle extremely natural and highlights their relevance for quantum cryptography. As outlined in Fig. 2, we imagine that an observer, Bob, can prepare an arbitrary state ρ_A which he will send to a referee, Alice. Alice then randomly chooses to perform one of two (or more) possible measurements, where we will use Θ to denote her choice of measurement. She records the outcome, K. She then tells Bob the choice of her measurement, i.e., she sends him Θ. Bob wins the game, if he correctly guesses Alice's measurement outcome K.

The uncertainty principle tells us that if Alice makes two incompatible measurements, then Bob cannot guess Alice's outcome for both measurements. This corresponds precisely to the notion of preparation uncertainty. It is indeed intuitive why such uncertainty relations form an important ingredient in proving the security of quantum cryptographic protocols, as we will explore in detail in Sec. VI. In the cryptographic setting ρ_A will be sent by an adversary trying to break a quantum cryptographic protocol. If Alice's measurements are incompatible, there is no way for the adversary to know the outcomes of both possible measurements with certainty - no matter what state he prepares.

The formulation of uncertainty relations as guessing games also makes it clear that there is an important twist to such games: What if Bob prepares a bipartite state ρ_{AB} and sends only the A part to Alice? That is, what if Bob's system is correlated with Alice's? Or, adopting the modern perspective of information, what if Bob has a non-trivial amount of *side information* about Alice's system? Traditional uncertainty relations implicitly assume that Bob has only classical side information. For

[3] Throughout this review, all logarithms are base 2.

[4] More precisely, the relation applies to non-degenerate observables on a finite-dimensional Hilbert space (see Sec. III.B).

example, he may remember a classical description of the state ρ_A or other details about the preparation. However, modern uncertainty relations—for example those derived by Berta *et al.* (2010) improving on work by Christandl and Winter (2005) and Renes and Boileau (2009)—allow Bob to have *quantum* rather than classical information about the state. As was already observed by Einstein *et al.* (1935), it is possible for Bob to experience no uncertainty at all in this case (in the sense that he can correctly guess Alice's measurement outcome K in the game described above).

We will devote Sec. IV to such modern uncertainty relations. It is these relations that will be of central importance in quantum cryptography, where the adversary may have gathered *quantum* and not just classical information during the course of the protocol that may reduce his uncertainty.

A. Scope of this review

Two survey articles partially discuss the topic of entropic uncertainty relations. Bialynicki-Birula and Rudnicki (2011) take a physics perspective and cover continuous variable entropic uncertainty relations and some discretized measurements. In contrast, Wehner and Winter (2010) take an information-theoretic perspective and discuss entropic uncertainty relations for discrete variables with an emphasis on relations that involve more than two measurements.

These reviews predate many recent advances in the field. For example, both reviews do not cover entropic uncertainty relations that take into account quantum correlations with the environment of the measured system. Moreover, applications of entropic uncertainty relations are only marginally discussed in both of these reviews. Here we discuss both physical and information-based applications. We therefore aim to give a comprehensive treatment of all of these topics in one reference, with the hope of benefiting some of the quickly emerging technologies that exploit quantum information.

There is an additional aspect of the uncertainty principle known as *measurement uncertainty*, see, e.g., Busch *et al.* (2007, 2014a); Hall (2004); and Ozawa (2003). The latter includes (1) joint measurability, the concept that there exists pairs of observables that cannot be measured simultaneously and (2) measurement–disturbance, the concept that there exist pairs of observables for which measuring one causes a disturbance of the other. Measurement uncertainty is still a highly debated topic of current research. We focus our review article on the concept of preparation uncertainty, although we briefly mention entropic approaches to measurement uncertainty in Sec. VII.C.

II. RELATION TO STANDARD DEVIATION APPROACH

Traditional formulations of the uncertainty principle, for example the ones due to Kennard and Robertson, measure uncertainty in terms of the standard deviation. In this section we argue why we think entropic formulations are preferable. For further discussion we refer to (Uffink, 1990).

A. Position and momentum uncertainty relations

For the case of position and momentum observables, the strength of the entropic formulation can be seen from the fact that the entropic uncertainty relation in (4) is strictly stronger, and in fact implies, the standard deviation relation (1). Following Białynicki-Birula and Mycielski (1975), we formally show that

$$h(Q) + h(P) \geq \log(e\pi) \quad \Longrightarrow \quad \sigma_Q \sigma_P \geq \frac{1}{2} \tag{6}$$

for all states, where here and henceforth in this article we work in units such that $\hbar = 1$. To see this, let us consider a random variable Q governed by a probability density $\Gamma(q)$ and the differential entropy

$$h(Q) = -\int_{-\infty}^{\infty} \Gamma(q) \log \Gamma(q) dq \,. \tag{7}$$

We assume that this quantity is finite in the following. Gaussian probability distributions,

$$\Gamma(q) = \frac{1}{\sqrt{2\pi\sigma_Q^2}} \exp\left(\frac{-(q-\overline{q})^2}{2\sigma_Q^2}\right), \tag{8}$$

where $\overline{q}$ denotes the mean, are special in the following sense: for a fixed standard deviation σ_Q, distributions of the form of (8) maximize the entropy in (7). It is a straightforward exercise to show this, e.g., using variational calculus with Lagrange multipliers.

It is furthermore straightforward to insert (8) into (7) to calculate the entropy of a Gaussian distribution

$$h(Q) = \log \sqrt{2\pi e \sigma_Q^2} \quad \text{(Gaussian)} \,. \tag{9}$$

Since Gaussians maximize the entropy, it follows that the following inequality holds in general

$$h(Q) \leq \log \sqrt{2\pi e \sigma_Q^2} \quad \text{(in general)} \,. \tag{10}$$

Now consider an arbitrary quantum state for a particle's translational degree freedom, which gives rise to random variables P and Q for the position and momentum respectively. Let us insert the resulting relations into (4) to find

$$\log(2\pi e \sigma_Q \sigma_P) = \log \sqrt{2\pi e \sigma_Q^2} + \log \sqrt{2\pi e \sigma_P^2} \tag{11}$$

$$\geq h(Q) + h(P) \tag{12}$$

$$\geq \log(e\pi) \,. \tag{13}$$

By comparing the left- and right-hand sides of (11), noting that the logarithm is a monotonic function, we see that (11) implies (1), and hence so does (4).

It is worth noting that (10) is a strict inequality if the distribution is non-Gaussian, and hence (4) is strictly stronger than (1) if the quantum state is non-Gaussian. While quantum mechanics textbooks often present (1) as *the* fundamental statement of the uncertainty principle, it is clear that (4) is stronger and yet not much more complicated. Furthermore, as discussed in Sec. IV the entropic formulation is more robust, allowing the relation to be easily generalized to situations involving correlations with the environment.

B. Finite spectrum uncertainty relations

As noted above, both the standard deviation and the entropy have been applied to formulate uncertainty relations for observables with a finite spectrum. However, it is largely unclear how the most popular formulations, Robertson's (2) and Maassen-Uffink's (5), are related. It remains an interesting open question whether there exists a formulation that unifies these two formulations. There is an important difference between (2) and (5) in that the former has a bound that depends on the state, while the latter only depends on the two observables.

Example 1. *Consider* (2) *for the case of a qubit, where* $X = |0\rangle\langle 1| + |1\rangle\langle 0|$ *and* $Z = |0\rangle\langle 0| - |1\rangle\langle 1|$*, corresponding to the* x*- and* z*-axes of the Bloch sphere. Then the commutator is proportional to the* Y *Pauli operator and the right-hand side of* (2) *reduces to* $(1/2)\langle Y\rangle$*. Hence* (2) *gives a trivial bound for all states that lie in the* xz *plane of the Bloch sphere. For the eigenstates of* X *and* Z*, this bound is tight since one of the two uncertainty terms is zero, and hence the trivial bound is a (perhaps undesirable) consequence of the fact that the left-hand side involves a* product *(rather than a sum) of uncertainties. However, for any other states in the* xz *plane, neither uncertainty is zero. This implies that* (2) *is* not tight *for these states.*

This example illustrates a major weakness of Robertson's relation for finite dimensional systems — it gives trivial bounds for certain states, even when the left-hand side is non-zero. This point was noted, e.g., by Deutsch (1983) and used as motivation for entropic uncertainty relations, which do not suffer from this weakness. However, we should note that Schrödinger (1930) slightly strengthened Robertson's bound by adding an additional state-dependent term that helps to get rid of the artificial trivial bound discussed in Ex. 1. Likewise, Maccone and Pati (2014) recently proved a state-dependent bound on the on the sum (not the product) of the two variances, and this bound also removes the trivial behavior of Robertson's bound. Furthermore, one still may be able to obtain a state-independent bound using standard deviation uncertainty measures in the finite-dimensional case. For example, Busch *et al.* (2014b) considered the qubit case and obtained a state-independent bound on the sum of the variances.

C. Advantages of entropic formulation

From a practical perspective, a crucial advantage of entropic uncertainty relations are their applications all throughout quantum cryptography. However, let us now mention several other reasons why the entropic formulation of the uncertainty principle is advantageous over the standard deviation formulation.

1. Counterintuitive behavior of standard deviation

It has been pointed out by several authors, for example by Bialynicki-Birula and Rudnicki (2011), that the standard deviation behaves somewhat strangely for some simple examples.

Example 2. *Consider a spin-1 particle with equal probability* $\Pr(s_z) = 1/3$ *to have each of the three possible values of* Z*-angular momentum,* $s_z \in \{-1, 0, 1\}$*. The standard deviation of the* Z*-angular momentum is* $\sigma_Z = \sqrt{2/3}$*. Now suppose we gain information about the spin such that we now know that it definitely does not have the value* $s_z = 0$*, and the new probability distribution is* $\Pr(1) = \Pr(-1) = 1/2$*,* $\Pr(0) = 0$*. We might expect the uncertainty to decrease, since we have gained information about the spin, but in fact the standard deviation increases, the new value being* $\sigma_Z = 1$*.*

We remark that the different behavior of standard deviation and entropy for spin angular momentum was recently highlighted by Dammeier *et al.* (2015), in the context of states that saturate the relevant uncertainty relation.

Bialynicki-Birula and Rudnicki (2011) noted an example for a particle's spatial position that is analogous to the above example.

Example 3. *Consider a very long box of length* L*, centered at* $x = 0$*, with two very small boxes of length* a *attached to the two ends of the long box, as depicted in Fig. 3. Suppose we know that a classical particle is confined to the two small end boxes, i.e., with equal probability it is one of the two small boxes. The standard deviation of the position is* $\sigma_x \approx L/2$*, assuming that* $L \gg a$*. Now suppose the barriers that separate the end boxes from the middle box are removed, and the particle is allowed to move freely between all three boxes. Intuitively one might expect that the uncertainty of the particle's position is now larger, since we now know nothing about where the*

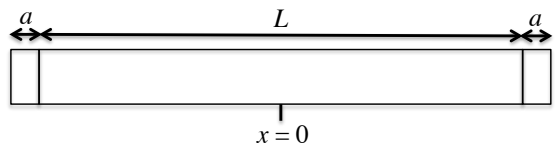


FIG. 3 Illustration for Ex. 3, where a particle is initially confined to the two small boxes at the end and excluded from the long middle box, and then the particle is allowed to go free into the middle box.

particle is inside the three boxes. However the new standard deviation is actually smaller: $\sigma_x \approx L/\sqrt{12}$.

Entropies on the other hand do not have this counterintuitive behavior, due to properties discussed below. Finally, let us note a somewhat obvious issue that, in some cases, a quantitative label (and hence the standard deviation) does not make sense, as illustrated in the following example.

Example 4. *Consider a neutrino's flavor, which is often modeled as a three-outcome observable with outcomes "electron", "muon", or "tau". As this is a non-quantitative observable, the standard deviation does not make sense in this context. Nevertheless, it is of interest to quantify the uncertainty about the neutrino flavor, i.e., how difficult it is to guess the flavor, which is naturally captured by the notion of entropy.*

2. Intuitive entropic properties

Deutsch (1983) emphasized that the standard deviation can change under a simple relabeling of the outcomes. For example, if one was to assign quantitative labels to the outcomes in Ex. 4, and then relabel them, the standard deviation would change. In contrast, the entropy is invariant under relabeling of outcomes, because it naturally captures the amount of *information* about a measurement outcome.

Furthermore, there is a nice monotonic property of entropy in the following sense. Suppose one does a random relabeling of the outcomes. One can think of this as a relabeling plus added noise, which naturally tends to spread the probability distribution out over the outcomes. Intuitively, a relabeling with the injection of randomness should never decrease the uncertainty. This property — non-decreasing under random relabeling — was highlighted by Friedland *et al.* (2013) as a desirable property of an uncertainty measure. Indeed, entropy satisfies this property. On the other hand, the physical process in Ex. 3 can be modeled mathematically as a random relabeling and hence we see the contrast in behavior between entropy and standard deviation.

Monotonicity under random relabeling is actually a special case of an even more powerful property. Think of the random relabeling as due to the fact that the observer is denied access to an auxiliary register that stores the information about which relabeling occurred. If the observer had access to the register, then their uncertainty would remain the same, but without access their uncertainty could potentially increase, but never decrease! More generally, this idea, that losing access to an auxiliary system cannot reduce one's uncertainty, is a desirable and powerful property of uncertainty measures known as the *data-processing inequality*. It is arguably a defining property of entropy measures, or more precisely, conditional entropy measures as discussed in Sec. IV.B. Furthermore this property is central in proving entropic uncertainty relations (Coles *et al.*, 2012).

3. Framework for classical and quantum systems

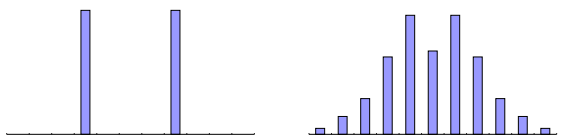
Entropy provides a robust mathematical framework that can be generalized to deal with correlated quantum systems. For example, the entropy framework allows us to discuss the uncertainty of an observable from the perspective of an observer who has access to part of the environment of the system, or to quantify quantum correlations like entanglement between two quantum systems. This requires measures of conditional uncertainty, namely conditional entropies. We highlight the utility of this framework in Sec. IV. A similar framework for standard deviation has not been developed.

4. Operational meaning and information applications

Perhaps the most compelling reason to consider entropy as the uncertainty measure of choice is that it has operational significance for various information-processing tasks. The standard deviation, in contrast, does not play a significant role in information theory. This is because entropy abstracts from the physical representation of information, as one can see from the following example.

Example 5. *Consider the two probability distributions in Fig. 4. They have the same standard deviation but very different entropy. The distribution in Fig. 4(a) has one bit of entropy since only two events are possible and occur with equal probability. If we want to record data from this random experiment this will require exactly one bit of storage per run. On the other hand, the distribution in Fig. 4(a) has approximately 3 bits of entropy and the recorded data cannot be compressed to less than 3 bits per run. Clearly, entropy has operational meaning in this context while standard deviation fails to distinguish these random experiments.*

Entropies have operational meaning for tasks such as randomness extraction (extracting perfect randomness



(a) low entropy distribution (b) high entropy distribution

FIG. 4 Two probability distributions with the same standard deviation but different entropy, as in Ex. 5.

from a partially random source) and data compression (sending minimal information to someone to help them guess the output of a partially random source). It is precisely these operational meanings that make entropic uncertainty relations useful for proving the security of quantum key distribution and other cryptographic tasks. We discuss such applications in Sec. VI.

The operational significance of entropy allows one to frame entropic uncertainty relations in terms of guessing games, see Sec. III.F and IV.D.1. These are simple yet insightful tasks where, e.g., one party is trying to guess the outcome of another party's measurements. Such games make it clear that the uncertainty principle is not just abstract mathematics; rather it is relevant to physical tasks that can be performed in a laboratory.

III. UNCERTAINTY WITHOUT A MEMORY SYSTEM

Historically, entropic uncertainty relations were first studied for position and momentum observables. However, to keep the discussion mathematically simple we begin here by introducing entropic uncertainty relations for finite-dimensional quantum systems, and we defer the discussion of infinite dimensions to Sec. V. It is worth noting that many physical systems of interest are finite-dimensional, such as photon polarization, neutrino flavor, and spin angular momentum.

In this section, we consider uncertainty relations where we measure a single preparation ρ_A. That is, there is no memory system. We emphasize that all uncertainty relations *with* a memory system can also be applied to the situation *without*. We will not list these special cases here, but refer to Section IV.

A. Entropy measures

Let us consider a discrete *random variable* X distributed according to the probability distribution P_X. We assume that X takes values in a finite set $\mathcal{X}$. For example, this set could be binary values $\{0,1\}$ or spin states $\{\uparrow,\downarrow\}$. In general, we will associate the random variable X with the outcome of a particular measurement. This random variable can take values $X = x$, where x is a specific instance of a measurement outcome that can be obtained with probability $P_X(X = x)$. However, entropies only depend on the probability law P_X and not on the specific labels of the elements in the set $\mathcal{X}$. Thus, we will in the following just assume this set to be of the form $[d] := \{1, 2, 3, \ldots, d\}$, where $d = |X|$ stands for the cardinality of the set $\mathcal{X}$.

1. Surprisal and Shannon entropy

Following Shannon (1948), we first define the *surprisal* of the event $X = x$ distributed according to P_X as $-\log P_X(x)$, often also referred to as *information content*. As its name suggests, the information content of $X = x$ gets larger when the event $X = x$ is less likely, i.e. when $P_X(x)$ is smaller. In particular, deterministic events have no information content at all, which is indeed intuitive since we learn nothing by observing an event that we are assured will happen with certainty. In contrast, the information content of very unlikely events can get arbitrarily large. Based on this intuition, the *Shannon entropy* is defined as

$$H(X) := \sum_x P_X(x) \log \frac{1}{P_X(x)}\,, \tag{14}$$

and quantifies the *average information content* of X. It is therefore a measure of the uncertainty of the outcome of the random experiment described by X. The Shannon entropy is by far the best-known measure of uncertainty and it is the one most commonly used to express uncertainty relations.

2. Rényi entropies

However, for some applications it is important to consider other measures of uncertainty that give more weight to events with high or low information content, respectively. For this purpose we employ a generalization of the Shannon entropy to a family of entropies introduced by Rényi (1961). The family includes several important special cases which we will discuss individually. These entropies have found many applications in cryptography and information theory (see Sec. VI), and have many convenient mathematical properties.[5]

[5] Another family of entropies that are often encountered are the Tsallis entropies (Tsallis, 1988). They have not found operational interpretation in cryptography or information theory, and we thus defer the discussion of Tsallis entropies until Sec. VII.A.

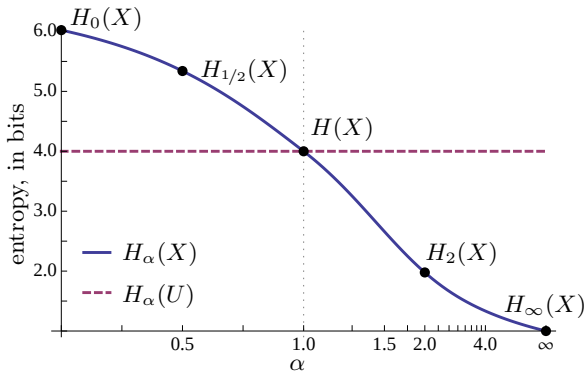


FIG. 5 Rényi entropies of X in Ex. 6 with $|X| = 65$ compared to a uniform random variable U on 4 bits.

The *Rényi entropy* of order α is defined as

$$H_\alpha(X) := \frac{1}{1-\alpha} \log \sum_x P_X(x)^\alpha \quad \text{for} \quad \alpha \in (0,1) \cup (1,\infty)\,, \tag{15}$$

and as the corresponding limit for $\alpha \in \{0, 1, \infty\}$. For $\alpha = 1$ the limit yields the Shannon entropy[6], and the Rényi entropies are thus a proper generalization of the Shannon entropy.

The Rényi entropies are monotonically decreasing as a function of α. Entropies with $\alpha > 1$ give more weight to events with high surprisal.

The *collision entropy*, $H_{\text{coll}} := H_2$, is given by

$$H_{\text{coll}}(X) = -\log p_{\text{coll}}(X)\,, \quad \text{where} \quad p_{\text{coll}}(X) := \sum_x P_X(x)^2 \tag{16}$$

is the collision probability, i.e. the probability that two independent instances of X are equal. The *min-entropy* $H_{\min} := H_\infty$, is of special significance in many applications. It characterizes the optimal probability of correctly guessing the value of X in the following sense

$$H_{\min}(X) = -\log p_{\text{guess}}(X)\,, \quad \text{where} \quad p_{\text{guess}}(X) := \max_x P_X(x)\,. \tag{17}$$

Clearly, the optimal guessing strategy is to bet on the most likely value of X, and the winning probability is then given by the maximum in (17). The min-entropy can also be seen as the minimum surprisal of X.

The Rényi entropies with $\alpha < 1$ give more weight to events with small surprisal. Noteworthy examples are the *max-entropy*, $H_{\max} := H_{1/2}$, and

$$H_0(X) = \log \big|\{x : P_X(x) > 0\}\big|\,, \tag{18}$$

where the latter is simply the logarithm of the essential support of P_X.

[6] It is a simple exercise to apply L'Hôpital's rule to (15) in the limit $\alpha \to 1$.

3. Examples and properties

For all the Rényi entropies, $H_\alpha(X) = 0$ if and only if the distribution is perfectly peaked, i.e., $P_X(x) = 1$ for some particular value x. On the other hand, the distribution $P_X(x) = |X|^{-1}$ is uniform if and only if the entropy takes its maximal value $H_\alpha(X) = \log |X|$.

The Rényi entropies can take on very different values depending on the parameter α as the following example visualized in Fig. 5 shows.

Example 6. *Note that the Rényi entropies for different values of α can be very different. To see this, consider a distribution of the form*

$$P_X(x) = \begin{cases} \frac{1}{2} & \text{for } x = 1 \\ \frac{1}{2(|X|-1)} & \text{else} \end{cases} \tag{19}$$

so that we have

$$H_{\min}(X) = \log 2\,, \quad \text{whereas} \quad H(X) = \log 2 + \frac{1}{2}\log(|X|-1) \tag{20}$$

is arbitrarily large as $|X| \geq 2$ increases. This is of particular relevance in cryptographic applications where $H_{\min}(X)$ — and not $H(X)$ — characterizes how difficult it is to guess a secret X. As we will see later, $H_{\min}(X)$ determines precisely the number of random bits that can be obtained from X.

Consider two probability distributions, P_X and Q_Y, and define $d = \max\{|X|, |Y|\}$. Now let us reorder the probabilities in P_X into a vector $P_X^\downarrow$ such that $P_X^\downarrow(1) \geq P_X^\downarrow(2) \geq \ldots \geq P_X^\downarrow(d)$, padding with zeros if necessary. Analogously arrange the probabilities in Q_Y into a vector $Q_Y^\downarrow$. We say P_X majorizes Q_Y and write $P_X \succ Q_Y$ if

$$\sum_{x=1}^{y} P_X^\downarrow(x) \geq \sum_{x=1}^{y} Q_Y^\downarrow(x), \quad \text{for all } y \in [d]\,. \tag{21}$$

Intuitively, the fact that P_X majorizes Q_Y means that P_X is less spread out than Q_Y. For example, the distribution $\{1, 0, \ldots, 0\}$ majorizes every other distribution, while the uniform distribution $\{|X|^{-1}, \ldots, |X|^{-1}\}$ is majorized by every other distribution.

One of the most fundamental properties of the Rényi entropy is that it is Schur-concave (Marshall *et al.*, 2011), meaning that it satisfies

$$H_\alpha(X) \leq H_\alpha(Y) \qquad \text{if } P_X \succ Q_Y\,. \tag{22}$$

This has an important consequence. Let $Y = f(X)$ for some (deterministic) function f. In other words, Y is obtained by *processing* X using the function f. The random variable Y is then governed by the push forward Q_Y of P_X, that is

$$Q_Y(y) = \sum_{x : f(x) = y} P_X(x)\,. \tag{23}$$

Clearly $P_X \prec Q_Y$ and thus we have $H_\alpha(X) \geq H_\alpha(Y)$. This corroborates our intuition that the input of a function is at least as uncertain as its output. If Z is just a reordering of X, or more generally if f is injective, then the two entropies are equal.

Finally we note that if two random variables X and Y are independent, we have

$$H_\alpha(XY) = H_\alpha(X) + H_\alpha(Y)\,. \tag{24}$$

This property is called *additivity*.

B. Preliminaries

1. Physical setup

The physical setup used throughout the remainder of this section is as follows. We consider a quantum system, A, that is measured in either one of two (or more) bases. The initial state of the system A is represented by a *density operator*, ρ_A, or more formally a positive semidefinite operator with unit trace acting on a finite-dimensional Hilbert space A. The measurements, for now, are given by two *orthonormal bases* of A. An orthonormal basis is a set of unit vectors in A that are mutually orthogonal and span the space A. The two bases are denoted by sets of rank-1 projectors,

$$\mathbb{X} = \left\{|\mathbb{X}^x\rangle\langle\mathbb{X}^x|\right\}_x \quad \text{and} \quad \mathbb{Z} = \left\{|\mathbb{Z}^z\rangle\langle\mathbb{Z}^z|\right\}_z\,. \tag{25}$$

We use projectors to keep the notation consistent as we will later consider more general measurements. This induces two random variables, X and Z corresponding to the measurement outcomes that result from measuring in the bases $\mathbb{X}$ and $\mathbb{Z}$, respectively. These are governed by the following probability laws, given by the Born rule. We have

$$P_X(x) = \langle\mathbb{X}^x|\rho_A|\mathbb{X}^x\rangle \quad \text{and} \quad P_Z(z) = \langle\mathbb{Z}^z|\rho_A|\mathbb{Z}^z\rangle\,, \tag{26}$$

respectively. We also note that $|X| = |Z| = d$, which is the dimension of the Hilbert space A.

2. Mutually unbiased bases

Before delving into uncertainty relations, let us remark that, in finite dimensions, there exist pairs of observables such that perfect knowledge about observable $\mathbb{X}$ implies complete ignorance about observable $\mathbb{Z}$. We say that such observables are unbiased, or mutually unbiased. For any finite dimensional space there exist pairs of orthonormal bases that satisfy this property. More precisely, two orthonormal bases $\mathbb{X}$ and $\mathbb{Z}$ are mutually unbiased bases (MUBs) if

$$|\langle\mathbb{X}^x|\mathbb{Z}^z\rangle|^2 = \frac{1}{d}, \quad \forall x, z\,. \tag{27}$$

In addition, a set of n orthonormal bases $\{\mathbb{X}_j\}$ is said to be a set of n MUBs if each basis $\mathbb{X}_j$ is mutually unbiased to every other basis $\mathbb{X}_k$, with $k \neq j$, in the set.

Example 7. *For a qubit the eigenvectors of the Pauli operators,*

$$\sigma_{\mathbb{X}} := |0\rangle\langle 1| + |1\rangle\langle 0| \tag{28}$$
$$\sigma_{\mathbb{Y}} := -i|0\rangle\langle 1| + i|1\rangle\langle 0| \tag{29}$$
$$\sigma_{\mathbb{Z}} := |0\rangle\langle 0| - |1\rangle\langle 1| \tag{30}$$

form a set of 3 MUBs.

In App. A we discuss constructions for sets of MUBs in higher dimensional spaces. We also point to Durt *et al.* (2010) for a review on this topic.

C. Measuring in two orthonormal bases

1. Shannon entropy

Based on the pioneering work by Deutsch (1983) and following a conjecture of Kraus (1987), Maassen and Uffink (1988) formulated entropic uncertainty relations for measurements of two complementary observables. Their best known relation uses the Shannon entropy to quantify uncertainty. It states that, for any state ρ_A,

$$H(X) + H(Z) \geq \log\frac{1}{c} =: q_{\mathrm{MU}}\,, \tag{31}$$

where the measure of incompatibility is a function of the *maximum overlap* of the two measurements, namely

$$c = \max_{x,z} c_{xz}, \quad \text{where } c_{xz} = |\langle\mathbb{X}^x|\mathbb{Z}^z\rangle|^2\,. \tag{32}$$

Note that q_{MU} is state-independent, namely independent of the initial state ρ_A. This is in contrast to Robertson's bound in (2).

The bound q_{MU} becomes non-trivial as long as $\mathbb{X}$ and $\mathbb{Z}$ do not have any vectors in common. Hence, in this case (31) shows that for any input density matrix there is some uncertainty in at least one of the two random variables X and Z, quantified by the Shannon entropies $H(X)$ and $H(Z)$, respectively. In general we have

$$\frac{1}{d} \leq c \leq 1 \quad \text{and hence} \quad 0 \leq q_{\mathrm{MU}} \leq \log d\,. \tag{33}$$

For the extreme case that $\mathbb{X}$ and $\mathbb{Z}$ are MUBs, as defined in (27), the overlap matrix $[c_{xz}]$ is flat: $c_{xz} = 1/d$ for all x and z, and the lower bound on the uncertainty then becomes maximal

$$H(X) + H(Z) \geq \log d\,. \tag{34}$$

Note that this is a necessary and sufficient condition: $c = 1/d$ iff the two bases are MUBs, and hence MUBs uniquely give the strongest uncertainty bound here.

For general observables $\mathbb{X}$ and $\mathbb{Z}$ the overlap matrix is not necessarily flat and the asymmetry of the matrix elements c_{xz} is quantified in (32) by taking the maximum over all x, z. In order to see why the maximum entry provides some (fairly coarse) measure of the flatness of the whole matrix, note that if the maximum entry of the overlap matrix is $1/d$, then all entries in the matrix must be $1/d$. Alternative measures of incompatibility will be discussed in Sec. III.C.5.

2. Rényi entropies

Maassen and Uffink (1988) also showed that the above relation (31) holds more generally in terms of Rényi entropies. For any $\alpha, \beta \geq \frac{1}{2}$ with $1/\alpha + 1/\beta = 2$, we have

$$H_\alpha(X) + H_\beta(Z) \geq q_{\mathrm{MU}} \,. \tag{35}$$

It is easily checked that the relation (31) in terms of the Shannon entropy is recovered for $\alpha = \beta = 1$. For $\alpha \to \infty$ with $\beta \to 1/2$ we get another interesting special case of (35) in terms of the min- and max-entropy

$$H_{\min}(X) + H_{\max}(Z) \geq q_{\mathrm{MU}} \,. \tag{36}$$

Since the min-entropy characterizes the probability of correctly guessing the outcome X, it is this type of relation that becomes most useful for applications in quantum cryptography and quantum information theory (see Sec. VI).

3. Proof of Maassen-Uffink

The original proof of (35) of Maassen and Uffink makes use of the Riesz-Thorin interpolation theorem. Recently an alternative proof was formulated (Coles *et al.*, 2012, 2011) using the monotonicity of the relative entropy under quantum channels. The latter approach is illustrated in App. B, where we prove the special case of the Shannon entropy relation (31). The proof is very simple and straightforward, hence we highly recommend the interested reader to see App. B. The Rényi entropy relation (35) follows from a more general proof given in App. C.3.

4. Tightness and extensions

Given the very simple and appealing form of the Maassen-Uffink relations (35) a natural question to ask is how tight these relations are. It is easily seen that if $\mathbb{X}$ and $\mathbb{Z}$ are MUBs, then they are tight for any of the states $\rho_A = |\mathbb{X}^x\rangle\langle\mathbb{X}^x|$ or $\rho_A = |\mathbb{Z}^z\rangle\langle\mathbb{Z}^z|$. Thus there cannot exist a better state-independent bound if $\mathbb{X}$ and $\mathbb{Z}$ are MUBs. However, for general orthonormal bases $\mathbb{X}$ and $\mathbb{Z}$ the relations (35) are not necessarily tight. This issue is addressed in the following subsections, where we also note that (31) can be tightened for mixed states ρ_A with a state-dependent bound.

Going beyond orthonormal bases, the above relations can be extended to more general measurements, as discussed in Sec. III.D. Finally, another interesting extension considers more than two observables (which in some cases leads to tighter bounds for two observables), as discussed in Sec. III.G.

5. Tighter bounds for qubits

Various attempts have been made to strengthen the Maassen–Uffink bound, particularly in the Shannon-entropy form (31). Let us begin by first discussing improvements upon (31) in the qubit case and then move on to arbitrary dimensions.

For $d = 2$ the situation is fairly simple since the overlap matrix $[c_{xz}]$ only depends on a single parameter, which we can take as the maximum overlap $c = \max_{x,z} c_{xz}$. Hence the goal is to find the largest function of c that still lower-bounds the entropic sum. Significant progress along these lines was made by Sánches-Ruiz (1998), who noted that the Maassen-Uffink bound, q_{MU}, could be replaced by a stronger bound

$$q_{\mathrm{SR}} = h_{\mathrm{bin}}\left(\frac{1+\sqrt{2c-1}}{2}\right). \tag{37}$$

Here, $h_{\mathrm{bin}}(p) := -p\log p - (1-p)\log(1-p)$ denotes the binary entropy.

Later work by Ghirardi *et al.* (2003) then attempted to find the optimal bound. They simplified the problem to a single-parameter optimization as follows

$$q_{\mathrm{opt}} = \min_\theta \left(h_{\mathrm{bin}}\left(\frac{1+\cos\theta}{2}\right) + h_{\mathrm{bin}}\left(\frac{1+\cos(\alpha-\theta)}{2}\right)\right) \tag{38}$$

where $\alpha := 2\arccos\sqrt{c}$. While it is straightforward to perform this optimization, Ghirardi *et al.* (2003) noted that an analytical solution could only be found for $c \gtrsim 0.7$. They found this analytical form to be

$$q_{\mathrm{G}} = 2h_{\mathrm{bin}}(b), \quad c \gtrsim 0.7\,, \tag{39}$$

$$\text{where } b := \left(\frac{1+\sqrt{c}}{2}\right). \tag{40}$$

Fig. 6 shows a plot of q_{opt}, q_{SR}, and q_{MU}. In addition, this plot also shows a bound q_{maj} obtained from a majorization technique discussed in Sec. III.I. For pairs of entropies H_α and H_β in (35), Abdelkhalek *et al.* (2015) recently completely characterized the amount of uncertainty in the one qubit case.

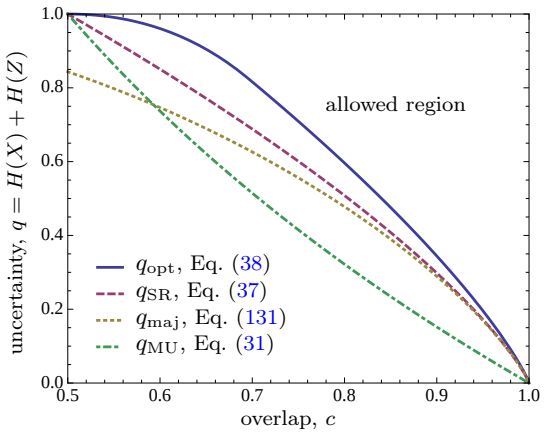


FIG. 6 Plot of various literature bounds on entropic uncertainty for qubit orthonormal bases, as a function of the maximum overlap c. The region above $q_{\rm opt}$ contains pairs (c, q) that can be achieved by quantum mechanics.

6. Tighter bounds in arbitrary dimension

Extending the qubit result above in (38), de Vicente and Sánchez-Ruiz (2008) found an analytical bound in the large overlap (i.e., large c) regime

$$q_{\rm dVSR} = 2h_{\rm bin}(b) \quad \text{for} \quad c \gtrsim 0.7\,, \tag{41}$$

which is stronger than the MU bound over this range, and they also obtained a numerical improvement over MU for the range $1/2 \leq c \lesssim 0.7$.

However the situation for $d > 2$ is more complicated than the qubit case. For $d > 2$ the overlap matrix $[c_{xz}]$ depends on more parameters than simply the maximum overlap c. Recent work has focused on exploiting these other overlaps to improve upon the MU bound. For example, Coles and Piani (2014b) derived a simple improvement on $q_{\rm MU}$ that captures the role of the second-largest entry of $[c_{xz}]$, denoted c_2, with the measure,

$$q_{\rm CP} = \log\frac{1}{c} + \frac{1}{2}(1-\sqrt{c})\log\frac{c}{c_2}\,. \tag{42}$$

Consider the following qutrit example where $q_{\rm CP} > q_{\rm MU}$.

Example 8. *Let $d = 3$ and consider the two orthonormal bases $\mathbb{X}$ and $\mathbb{Z}$ related by the unitary transformation,*

$$U = \begin{pmatrix} 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \\ 1/\sqrt{2} & 0 & -1/\sqrt{2} \\ 1/\sqrt{6} & -\sqrt{2/3} & 1/\sqrt{6} \end{pmatrix}\,. \tag{43}$$

We have $q_{\rm MU} = \log(3/2) \approx 0.58$ while $q_{\rm CP} \approx 0.64$.

More recently, a bound similar in spirit to $q_{\rm CP}$ was obtained by Rudnicki *et al.* (2014), of the form

$$q_{\rm RPZ} = \log\frac{1}{c} - \log\left(b^2 + \frac{c_2}{c}(1-b^2)\right)\,. \tag{44}$$

Note that $q_{\rm RPZ} \geq q_{\rm MU}$. However there is no clear relation between $q_{\rm CP}$ and $q_{\rm RPZ}$.

For arbitrary pairs of entropies H_α and H_β, Abdelkhalek *et al.* (2015) give conditions on the minimizing state of (35). In particular, the minimizing state is pure and real. For measurements in the standard and Fourier basis, further conditions are obtained.

7. Tighter bounds for mixed states

Notice that (31) can be quite loose for mixed states. For example, if $\rho_A = \mathbb{1}/d$, then the left-hand side of (31) is $2\log d$, whereas the right-hand side is at most $\log d$. This looseness can be addressed by introducing a state-dependent bound that gets larger as ρ_A becomes more mixed. The mixedness of ρ_A can be quantified by the von Neumann entropy $H(\rho_A)$, which we also denote by $H(A)_\rho$, defined by

$$H(\rho_A) = -\operatorname{tr}\left[\rho_A \log \rho_A\right] = \sum_j \lambda_j \log\frac{1}{\lambda_j}\,, \tag{45}$$

where an eigenvalue decomposition of the state is given by $\rho_A = \sum_j \lambda_j |\phi_j\rangle\langle\phi_j|_A$. Note that $0 \leq H(\rho_A) \leq \log d$, where $H(\rho_A) = 0$ for pure states and $H(\rho_A) = \log d$ for maximally mixed states. In the literature, the von Neumann entropy is sometimes also denoted using $S(A) = H(A)$. However, here we will follow the more common convention in quantum information theory. We note that the entropy never decreases when applying a measurement $\mathbb{X} = \{|\mathbb{X}^x\rangle\langle\mathbb{X}^x|\}_x$ to ρ_A, that is,

$$H(\rho_A) \leq H(X)_P \quad \text{with} \quad P_X(x) = \langle\mathbb{X}^x|\rho_A|\mathbb{X}^x\rangle\,. \tag{46}$$

Equation (31) was strengthened for mixed states by Berta *et al.* (2010), with the bound

$$H(X) + H(Z) \geq \log\frac{1}{c} + H(\rho_A)\,. \tag{47}$$

A proof of (47) is given in App. B; see also Frank and Lieb (2012) for a direct matrix analysis proof. When $\mathbb{X}$ and $\mathbb{Z}$ are MUBs, this bound is tight for any state ρ_A that is diagonal in either the $\mathbb{X}$ or $\mathbb{Z}$ basis.

D. Arbitrary measurements

Many interesting measurements are not of the orthonormal basis form. For example, coarse-grained (degenerate) projective measurements are relevant to probing macroscopic systems. Also, there are other measurements that are informationally complete in the sense that their statistics allow one to reconstruct the density operator.

The most general description of measurements in quantum mechanics is that of positive operator valued measures (POVMs). A POVM on a system A is a set of

positive semidefinite operators $\{\mathbb{X}^x\}$ that decompose the identity, $\sum_x \mathbb{X}^x = \mathbb{1}_A$. The number of POVM elements in the set can be much larger or much smaller than the Hilbert space dimension of the system. Physically, a POVM can be implemented as a joint coarse-grained measurement on the system of interest with an ancilla system.

For two POVMs $\mathbb{X} = \{\mathbb{X}^x\}_x$ and $\mathbb{Z} = \{\mathbb{Z}^z\}_z$, the general Born rule now induces the distributions

$$P_X(x) = \mathrm{tr}\left[\rho_A \mathbb{X}^x\right] \quad \text{and} \quad P_Z(z) = \mathrm{tr}\left[\rho_A \mathbb{Z}^z\right]. \tag{48}$$

Krishna and Parthasarathy (2002) proposed an incompatibility measure using the operator norm. Namely they considered

$$c = \max_{x,z} c_{xz} \quad \text{with} \quad c_{xz} = \left\| \sqrt{\mathbb{X}^x}\sqrt{\mathbb{Z}^z} \right\|^2, \tag{49}$$

where $\|\cdot\|$ denotes the operator norm (i.e. the maximal singular value). Using this measure they generalized (31) to the case of POVMs, i.e., we still have

$$H(X) + H(Z) \geq \log \frac{1}{c}, \tag{50}$$

but now using the generalized version of c in (49). More recently, Tomamichel (2012) noted that an alternative generalization to POVMs is obtained by replacing c with

$$c' := \min\left\{ \max_x \left\| \sum_z \mathbb{Z}^z \mathbb{X}^x \mathbb{Z}^z \right\|, \max_z \left\| \sum_x \mathbb{X}^x \mathbb{Z}^z \mathbb{X}^x \right\| \right\}, \tag{51}$$

and the author conjectured that c' always provides a stronger bound than c.

Example 9. *Consider two POVMs given by*

$$\mathbb{X} = \mathbb{Z} = \frac{1}{2}\Big\{ |0\rangle\langle 0|, |1\rangle\langle 1|, |+\rangle\langle +|, |-\rangle\langle -| \Big\}. \tag{52}$$

For these POVMs we find $c = 1/4$, but $c' = 3/16$ is strictly smaller.

Indeed this conjecture was proved by Coles and Piani (2014b), where it was shown that

$$\left\| \sum_z \mathbb{Z}^z \mathbb{X}^x \mathbb{Z}^z \right\| \leq \max_z c_{xz}, \tag{53}$$

and hence $c' \leq c$, implying that $\log(1/c')$ provides a stronger bound on entropic uncertainty than $\log(1/c)$.

Interestingly, a general POVM can have a non-trivial uncertainty relation on its own. That is, for some POVM $\mathbb{X}$, there may not exist any state ρ_A that has $H(X) = 0$. Krishna and Parthasarathy (2002) noted this and derived the single POVM uncertainty relation

$$H(X) \geq -\log \max_x \left\| \mathbb{X}^x \right\|. \tag{54}$$

In fact the proof is straightforward: simply apply (50) to the case where $\mathbb{Z} = \{\mathbb{1}\}$ is the trivial POVM. The relation (54) can be further strengthened by apply this approach to c' in (51), instead of c.

E. State-dependent measures of incompatibility

In most uncertainty relations we have encountered so far, the measure of incompatibility, for example the overlap c, is a function of the measurements employed but is independent of the quantum state prior to measurement. The sole exception is the strengthened Maassen-Uffink relation in (47) where the lower bound is the sum of an ordinary, state-independent measure of incompatibility and the entropy of the pre-measurement state. In the following, we review some uncertainty relations that use measures of incompatibility that are *state dependent*.

It was shown in (Tomamichel and Hänggi, 2013) that the Maassen-Uffink relation (31) also holds when the overlap c is replaced by an effective overlap, denoted c^*. Informally, c^* is given by the average overlap of the two measurements on different subspaces of the Hilbert space, averaged using the probability of finding the state in the subspace. We refer the reader to the paper mentioned above for a formal definition of c^*. Here we discuss a simple example showing that state-dependent uncertainty relations can be significantly tighter.

Example 10. *Let us apply one out of two projective measurements, either in the orthonormal basis*

$$\Big\{|0\rangle, |1\rangle, |\perp\rangle\Big\} \quad \text{or} \quad \Big\{|+\rangle, |-\rangle, |\perp\rangle\Big\} \tag{55}$$

on a state ρ which has the property that '$\perp$' is measured with probability at most ε.[7] *The Maassen-Uffink relation (31) gives a trivial bound as the overlap of the two bases is $c = 1$ due to the vector $|\perp\rangle$ that appears in both bases. Still, our intuitive understanding is that the uncertainty about the measurement outcome is high as long as ε is small. In fact, the* effective overlap *(Tomamichel and Hänggi, 2013) of this setup is*

$$c^* = (1-\varepsilon)\frac{1}{2} + \varepsilon, \tag{56}$$

and captures this intuition. This formula can be interpreted as follows: with probability $1-\varepsilon$ we are in the subspace spanned by $|0\rangle$ and $|1\rangle$, where the overlap is $1/2$, and with probability ε we measure $\perp$ and have full overlap.

An alternative approach to state-dependent uncertainty relations was introduced in (Coles and Piani, 2014b). They showed that the factor $q_{\mathrm{MU}} = \log(1/c)$ in the Maassen-Uffink relation (31) can be replaced by the state-dependent factor

$$q(\rho_A) := \max\{q_X(\rho_A), q_Z(\rho_A)\}, \quad \text{where} \tag{57}$$

$$q_X(\rho_A) := \sum_x P_X(x) \log \frac{1}{\max_z c_{xz}}, \tag{58}$$

[7] The diagonal states are $|\pm\rangle = (|0\rangle \pm |1\rangle)/\sqrt{2}$.

and $q_Z(\rho_A)$ is defined analogously to $q_X(\rho_A)$, but with x and z interchanged. Here, $P_X(x)$ and c_{xz} are given by (26) and (32), respectively. Note that this strengthens the Maassen-Uffink bound, $q(\rho_A) \geq q_{\mathrm{MU}}$, since averaging $\log(1/\max_z c_{xz})$ over all x is larger than minimizing it over all x. In many cases $q(\rho_A)$ is significantly stronger than q_{MU} as discussed in (Coles and Piani, 2014b).

Recently Kaniewski *et al.* (2014) derived entropic uncertainty relations in terms of the *effective anti-commutator* of arbitrary binary POVMs $\mathbb{X} = \{\mathbb{X}^0, \mathbb{X}^1\}$ and $\mathbb{Z} = \{\mathbb{Z}^0, \mathbb{Z}^1\}$, namely the quantity

$$\varepsilon^* = \frac{1}{2}\operatorname{tr}\left[\rho[O_{\mathbb{X}}, O_{\mathbb{Z}}]_+\right] = \frac{1}{2}\operatorname{tr}\left[\rho(O_{\mathbb{X}}O_{\mathbb{Z}} + O_{\mathbb{Z}}O_{\mathbb{X}})\right], \quad \text{where} \quad O_{\mathbb{X}} = \mathbb{X}^0 - \mathbb{X}^1 \quad \text{and} \quad O_{\mathbb{Z}} = \mathbb{Z}^0 - \mathbb{Z}^1 \tag{59}$$

are binary observables corresponding to the POVMs $\mathbb{X}$ and $\mathbb{Z}$, respectively. In (59), we use the notation $[\cdot,\cdot]_+$ to denote the anti-commutator. We note that $\varepsilon^* \in [-1, 1]$. This results, for example, in the following uncertainty relation for the Shannon entropy:

$$H(X) + H(Z) \geq h_{\mathrm{bin}}\left(\frac{1 + \sqrt{|\varepsilon^*|}}{2}\right). \tag{60}$$

We refer the reader to (Kaniewski *et al.*, 2014) for similar uncertainty relations in terms of Rényi entropies as well as extensions to more than two measurements. Finally, for measurements acting on qubits, we find that $|\varepsilon^*| = 2c - 1$, and (60) hence reduces to the Sanchez-Ruiz bound (37).

F. Relation to guessing games

Let us now explain in detail how some of the relations above can be interpreted in terms of a guessing game. We elaborate on the brief discussion of guessing games in Sec. I, and we refer the reader back to Fig. 2 for an illustration of the game.

The game is as follows. Suppose that Bob prepares system A in state ρ_A. He then sends A to Alice, who randomly performs either the $\mathbb{X}$ or $\mathbb{Z}$ measurement. The measurement outcome is a bit denoted as K, and Bob's task is to guess K, given that he received the basis choice denoted by $\Theta \in \{\theta_{\mathbb{X}}, \theta_{\mathbb{Z}}\}$ from Alice.

We can rewrite the Maassen-Uffink relation (31) in the following way such that the connection to the above guessing game becomes transparent. Denoting the standard basis on A as $\{|k\rangle\}_{k=1}^d$, we write

$$|\mathbb{X}^k\rangle =: U_{\mathbb{X}}|k\rangle \quad \text{and} \quad |\mathbb{Z}^k\rangle =: U_{\mathbb{Z}}|k\rangle. \tag{61}$$

Then, we have

$$\frac{1}{2}\Big(H(K|\Theta = \theta_{\mathbb{X}}) + H(K|\Theta = \theta_{\mathbb{Z}})\Big) \geq \frac{1}{2}\log\frac{1}{c}, \tag{62}$$

with the conditional probability distribution

$$P_{K|\Theta=\theta_{\mathbb{X}}}(k) := \langle k|U_{\mathbb{X}}^\dagger \rho U_{\mathbb{X}}|k\rangle \quad \text{for} \quad k = 1, \ldots, d \tag{63}$$

and similarly for $\theta_{\mathbb{Z}}$. Alternatively we can also write this as

$$H(K|\Theta) \geq \frac{1}{2}\log\frac{1}{c} \quad \text{with} \quad \Theta \in \{\theta_{\mathbb{X}}, \theta_{\mathbb{Z}}\}, \tag{64}$$

in terms of the conditional Shannon entropy

$$H(K|\Theta) := H(K\Theta) - H(\Theta) \tag{65}$$

$$= \frac{1}{2}\Big(H(K|\Theta = \theta_{\mathbb{X}}) + H(K|\Theta = \theta_{\mathbb{X}})\Big) \tag{66}$$

of the bipartite distribution

$$P_{K\Theta}(k, \theta_j) := \frac{1}{2}\langle k|U_j^\dagger \rho U_j|k\rangle \quad \text{with} \quad k = 1, \ldots, d \quad j = \mathbb{X}, \mathbb{Z}. \tag{67}$$

That is, each measurement labeled θ_j is chosen with equal probability 1/2 and we condition on this choice. Notice that the form in (64) is connected to the guessing game in Fig. 2. Regardless of the state ρ_A that Bob prepares, the uncertainty relation (64) implies that he will not be able to win the guessing game (i.e., perfectly guess K) if $c < 1$. Bob might be able to partially guess K, but never with perfect accuracy. In this sense, the Maassen-Uffink relation *is a fundamental constraint on one's ability to win a guessing game.*

Actually, in the context of guessing games, the min-entropy is the more operational measure than the Shannon entropy. For example, a diligent reading of Deutsch (1983) reveals the relation

$$p_{\mathrm{guess}}(X) \cdot p_{\mathrm{guess}}(Z) \leq b^2, \tag{68}$$

for orthonormal bases $\mathbb{X}$ and $\mathbb{Z}$, where b is defined in (40). This relation gives an upper bound on the product of the guessing probabilities (or equivalently, a lower bound on the sum of the min-entropies) associated with X and Z. However, to make a more explicit connection to the guessing game described above, one would like to upper bound the sum (or average) of the guessing probabilities, namely the quantity

$$p_{\mathrm{guess}}(K|\Theta) = \frac{1}{2}\Big(p_{\mathrm{guess}}(K|\Theta = \theta_{\mathbb{X}}) + p_{\mathrm{guess}}(K|\Theta = \theta_{\mathbb{Z}})\Big). \tag{69}$$

Indeed, the quantity (69) can easily be upper-bounded Schaffner (2007) as

$$p_{\mathrm{guess}}(K|\Theta) \leq b \quad \text{or equivalently} \tag{70}$$

$$H_{\min}(K|\Theta) \geq \log\frac{1}{b}. \tag{71}$$

Example 11. *For the Pauli qubit measurements* $\{\sigma_{\mathbb{X}}, \sigma_{\mathbb{Z}}\}$ *the min-entropy uncertainty relation* (71) *becomes*

$$H_{\min}(K|\Theta) \geq \log \frac{2\sqrt{2}}{1+\sqrt{2}} \,. \tag{72}$$

We emphasize that $p_{\text{guess}}(K|\Theta)$ is precisely the probability for winning the game described in Fig. 2. Hence, the entropic uncertainty relation (71) gives *the* fundamental limit on winning the game. Finally, we remark that (71) is stronger than Deutsch's relation (68), due to the following argument. For the min-entropy, conditioning on the measurement choice is defined as

$$\begin{aligned} H_{\min}(K|\Theta) &:= -\log\left(\frac{1}{2}\sum_{j=1,2} 2^{-H_{\min}(K|\Theta=\theta_j)}\right) \\ &\neq H_{\min}(K\Theta) - H_{\min}(\Theta) \quad \text{(in general)}\,, \end{aligned} \tag{73}$$

in contrast to the Shannon entropy in (65). However in analogy to (66), we have

$$H_{\min}(K|\Theta) \leq \frac{1}{2}\sum_{j=1,2} H_{\min}(K|\Theta=\theta_j)\,. \tag{74}$$

due to the concavity of the logarithm. For a general discussion of conditional entropies we point to Sec. IV.B.

G. Multiple measurements

So far we have only considered entropic uncertainty relations quantifying the complementarity of two measurements. However, there is no fundamental reason for restricting to this setup and in the following we discuss entropic uncertainty relations for multiple measurements. We mostly focus on special sets of measurements that generate strong uncertainty relations. This is of particular interest for various applications in quantum cryptography (see Sec. VI.C).

The notation introduced for guessing games in Sec. III.F is particularly useful in the multiple measurements setting. In this notation, for larger sets of measurements we are interested in finding lower bounds of the form

$$H(K|\Theta) \geq f(\Theta, \rho_A) > 0 \;\text{ with }\; \Theta \in \{\theta_1, \ldots, \theta_L\}\,, \tag{75}$$

where, similarly to (67),

$$P_{K\Theta}(k,\theta_j) := \frac{1}{L}\langle k|U_j^\dagger \rho U_j|k\rangle \;\text{ with }\; \begin{aligned} k &= 1,\ldots,d \\ j &= 1,\ldots,L\,. \end{aligned} \tag{76}$$

Again the left-hand side of (75) might alternatively be written as

$$H(K|\Theta) = \frac{1}{L}\sum_{j=1}^{L} H(K|\Theta=\theta_j)\,, \tag{77}$$

where the conditional probability distribution $P_{K|\Theta=\theta_j}$ is defined analogously to (63).

1. Bounds implied by two measurements

It is important to realize that, e.g., the Maassen-Uffink relation (31) already implies bounds for larger sets of measurements. This is easily seen by just taking the Maassen-Uffink relation for all possible pairs of measurements and adding the corresponding lower bounds.

Example 12. *For the qubit Pauli measurements we find by the simple iterative application of the tightened Maassen-Uffink bound* (47) *for the measurement pairs* $\{\sigma_{\mathbb{X}}, \sigma_{\mathbb{Y}}\}$, $\{\sigma_{\mathbb{X}}, \sigma_{\mathbb{Z}}\}$, *and* $\{\sigma_{\mathbb{Y}}, \sigma_{\mathbb{Z}}\}$ *that*

$$H(K|\Theta) \geq \frac{1}{2} + \frac{1}{2}H(\rho_A) \;\text{ with }\; \Theta \in \{\sigma_{\mathbb{X}}, \sigma_{\mathbb{Y}}, \sigma_{\mathbb{Z}}\}\,. \tag{78}$$

The goal of this section is to find uncertainty relations that are stronger than any bounds that can be derived directly from relations for two measurements.

2. Mutually unbiased bases

At first glance, one might think that measuring in mutually unbiased bases always results in a large amount of uncertainty. Somewhat surprisingly, this is not the case. In fact, Ballester and Wehner (2007) show that for $d = p^{2l}$ with p prime and $l \in \mathbb{N}$, there exist up to $L = p^l + 1$ many MUBs together with a state ρ_A for which

$$H(K|\Theta) = \frac{\log d}{2} \;\text{ with }\; \Theta \in \{\theta_1, \ldots, \theta_L\}\,. \tag{79}$$

That is, we observe no more uncertainty than if we had just considered two incompatible measurements. While a certain amount of mutual unbiasedness is therefore a necessary condition for strong uncertainty relations, it is in general not sufficient.

Starting from the two measurements setting, a promising candidate for deriving strong uncertainty relations are sets of MUBs (see Sec. III.B.2 for an introduction and Sec. A for details). For a full set of $d+1$ MUBs it was shown that (Ivanovic, 1992; Larsen, 1990; Sánchez, 1993),

$$H(K|\Theta) \geq \log(d+1) - 1 \;\text{ with }\; \Theta \in \{\theta_1, \ldots, \theta_{d+1}\}\,. \tag{80}$$

This is a very strong bound since the entropic term on the left-hand side can become at most $\log d$ for any number and choice of measurements. The relation (80) can be derived from an uncertainty equality for the collision entropy H_{coll}. Namely, we have for any quantum state ρ_A on a d-dimensional system and a full set of

$d+1$ MUBs (Ballester and Wehner, 2007; Brukner and Zeilinger, 1999; Ivanovic, 1992),

$$H_{\mathrm{coll}}(K|\Theta) = \log(d+1) - \log\left(2^{-H_{\mathrm{coll}}(\rho_A)}+1\right) \quad \text{with } \Theta \in \{\theta_1,\dots,\theta_{d+1}\}\,, \tag{81}$$

where for the collision entropy the conditioning on the measurement choice is defined as

$$H_{\mathrm{coll}}(K|\Theta) := -\log\left(\frac{1}{L}\sum_{j=1}^{L} 2^{-H_{\mathrm{coll}}(K|\Theta=\theta_j)}\right) \neq H_{\mathrm{coll}}(K\Theta) - H_{\mathrm{coll}}(\Theta) \quad \text{(in general)}\,. \tag{82}$$

We also refer to Sec. III.A for a general discussion about conditional entropies. Moreover, the quantum collision entropy is a measure for how mixed the state ρ_A is and defined as

$$H_{\mathrm{coll}}(\rho_A) := -\log \mathrm{tr}\left[\rho_A^2\right]\,. \tag{83}$$

We emphasize that since (81) is an equality it is tight for every state. By the concavity of the logarithm we also have in analogy to the Shannon entropy (77),

$$H_{\mathrm{coll}}(K|\Theta) \leq \frac{1}{d+1}\sum_{j=1}^{d+1} H_{\mathrm{coll}}(K|\Theta=\theta_j)\,. \tag{84}$$

Example 13. *For the qubit Pauli measurements (81) yields*

$$H_{\mathrm{coll}}(K|\Theta) \geq \log 3 - \log\left(2^{-H_{\mathrm{coll}}(\rho_A)}+1\right) \quad \textit{with } \Theta \in \{\sigma_{\mathbb{X}},\sigma_{\mathbb{Y}},\sigma_{\mathbb{Z}}\}\,. \tag{85}$$

The uncertainty relation for the Shannon entropy (80) follows from (81) by at first only considering pure states, i.e., states with $H_{\mathrm{coll}}(\rho_A)=0$, and using that the Rényi entropies are monotonically decreasing as a function of the parameter α (note that the collision entropy corresponds to $\alpha=2$ and the Shannon entropy to $\alpha=1$). For mixed states ρ_A we can extend this in a second step by taking the eigendecomposition and making use of the concavity of the Shannon entropy. For later purposes we note that it is technically often more accessible to work with the collision entropy H_{coll} (even when ultimately interested in uncertainty relations in terms of other entropies).

The uncertainty relation (80) was improved for d even to (Sánchez-Ruiz, 1995),

$$H(K|\Theta) \geq \frac{1}{d+1}\left(\frac{d}{2}\log\left(\frac{d}{2}\right) + \left(\frac{d}{2}+1\right)\log\left(\frac{d}{2}+1\right)\right) \quad \text{with } \Theta \in \{\theta_1,\dots,\theta_{d+1}\}\,. \tag{86}$$

Example 14. *For the qubit Pauli measurements (86) yields*

$$H(K|\Theta) \geq \frac{2}{3} \quad \text{with } \Theta \in \{\sigma_{\mathbb{X}},\sigma_{\mathbb{Y}},\sigma_{\mathbb{Z}}\}\,, \tag{87}$$

also see (Sánches-Ruiz, 1998) for more on this special case. Moreover, with (Coles et al., 2011) we can add an entropy dependent term on the right-hand side,

$$H(K|\Theta) \geq \frac{2}{3} + \frac{1}{3}H(\rho_A) \quad \text{with } \Theta \in \{\sigma_{\mathbb{X}},\sigma_{\mathbb{Y}},\sigma_{\mathbb{Z}}\}\,. \tag{88}$$

Note that is never a worse bound than (78) which just followed from the tightened Maassen-Uffink relation for two measurements (47). Moreover, the relation (87) becomes an equality for any eigenstate of the Pauli measurements, while (88) becomes an equality for any state ρ_A that is diagonal in the eigenbasis of one of the Pauli measurements.

For smaller sets of $L<d+1$ MUBs we immediately get a weak bound from an iterative application of the Maassen-Uffink relation (31) for MUBs,

$$H(K|\Theta) \geq \frac{\log d}{2} \quad \text{with } \Theta \in \{\theta_1,\dots,\theta_L\}\,. \tag{89}$$

However, it surprisingly turns out that the bound (89) can in general not be improved much.

Example 15. *In $d=3$ there exists a set of $L=3$ MUBs together with a state ρ_A such that (Wehner and Winter, 2010),*

$$H(K|\Theta) = 1 \quad \text{with } \Theta \in \{\theta_1,\theta_2,\theta_3\}\,. \tag{90}$$

This only allows a relatively weak uncertainty relation for $L=d=3$ and we have from (Wu et al., 2009),

$$H(K|\Theta) \geq \frac{8}{9} \approx 0.89 \quad \text{with } \Theta \in \{\theta_1,\theta_2,\theta_3\}\,. \tag{91}$$

This is slightly stronger than the lower bound from (89),

$$H(K|\Theta) \geq \frac{\log 3}{2} \approx 0.79 \quad \text{with } \Theta \in \{\theta_1,\theta_2,\theta_3\}\,. \tag{92}$$

Again, in general this only allows relatively weak uncertainty relations for $L<d+1$ and we have from (Wu *et al.*, 2009),

$$H_{\mathrm{coll}}(K|\Theta) \geq -\log\frac{d\cdot 2^{-H_{\mathrm{coll}}(\rho_A)}+L-1}{L\cdot d} \quad \text{with } \Theta \in \{\theta_1,\dots,\theta_L\}\,. \tag{93}$$

This implies in particular the Shannon entropy relation (Azarchs, 2004),

$$H(K|\Theta) \geq -\log\frac{d+L-1}{L\cdot d} \quad \text{with } \Theta \in \{\theta_1,\dots,\theta_L\}\,, \tag{94}$$

see also (Wehner and Winter, 2010) for an elementary proof. For comparison, with $L = d = 3$, (94) yields

$$H(K|\Theta) \geq \log \frac{9}{5} \approx 0.85 \quad \text{with} \quad \Theta \in \{\theta_1, \theta_2, \theta_3\}\,, \tag{95}$$

which is between (89) and (91). Additional evidence that general sets of less than $d+1$ MUBs in dimension d only generate weak uncertainty relations is presented by (Ambainis, 2010; Ballester and Wehner, 2007; DiVincenzo *et al.*, 2004; Hasegawa *et al.*, 2007). Many of the findings here also extend to the setting of approximate mutually unbiased bases (Hayden *et al.*, 2004).

In terms of the min-entropy, Mandayam *et al.* (2010) show that for measurements in L possible MUBs the following two bounds hold

$$\frac{1}{L}\sum_{\theta=1}^{L} H_{\min}(K|\Theta=\theta) \geq -\log\left[\frac{1}{d}\left(1+\frac{d-1}{\sqrt{L}}\right)\right]\,, \tag{96}$$

$$\frac{1}{L}\sum_{\theta=1}^{L} H_{\min}(K|\Theta=\theta) \geq -\log\left[\frac{1}{L}\left(1+\frac{L-1}{\sqrt{d}}\right)\right]\,. \tag{97}$$

Each of these is better in certain regimes, and the latter can indeed be tight. They also study uncertainty relations for certain classes of MUBs that exhibit special symmetry properties. Furthermore, uncertainty relations for a full set of $L = d+1$ MUBs can also be expressed in terms of the extrema of Wigner functions (Mandayam *et al.*, 2010; Wootters and Sussman, 2007). It remains an interesting topic to study uncertainty relations for MUBs and in Sec. III.G.7 we present some related results of Kalev and Gour (2014).

3. Measurements in random bases

Another candidate for strong uncertainty relations are sets of measurements that are chosen at random. Extending on the previous results of Hayden *et al.* (2004), Fawzi *et al.* (2011) show that in dimension d and for any number $L > 2$ of measurements chosen at random,[8] there exists a constant C such that,

$$H(K|\Theta) \geq \log d \cdot \left(1 - \sqrt{\frac{1}{L}\cdot C\log(L)}\right) - g(L) \quad \text{with} \quad \Theta \in \{\theta_1, \ldots, \theta_L\}\,, \tag{98}$$

[8] By "at random" we mean according to the Haar measure on the unitary group.

and the fudge term $g(L) = O\left(\log\left(L/\log(L)\right)\right)$. Note that for any set of L measurements we always have

$$H(K|\Theta) \leq \log d \cdot \left(1 - \frac{1}{L}\right) \quad \text{with} \quad \Theta \in \{\theta_1, \ldots, \theta_L\}\,. \tag{99}$$

Hence, the relation (98) is already reasonably strong. However, very recently (98) was improved by proving a conjecture as stated in (Wehner and Winter, 2010). Namely, Adamczak *et al.* (2014) show that in dimension d and for any number $L > 2$ measurements chosen at random, there exists a constant D such that,

$$H(K|\Theta) \geq \log d \cdot \left(1 - \frac{1}{L}\right) - D \quad \text{with} \quad \Theta \in \{\theta_1, \ldots, \theta_L\}\,. \tag{100}$$

We emphasize that this matches the upper bound (99) up to the constant D.

The downside with the relations (98) and (100), however, is that measurements are not explicit. This is an issue for applications. In particular, it is computationally inefficient to sample from the Haar measure. Fawzi *et al.* (2011) show that the measurements in their relation (98) can be made explicit and efficient if the number L of measurements is small enough. More precisely, for n qubits (with n sufficiently large) and $\varepsilon > 0$, there exists a constant C and a set of

$$L \leq (n/\varepsilon)^{C\log(1/\varepsilon)} \tag{101}$$

measurements generated by unitaries computable by quantum circuits of size $O(\text{polylog}n)$ such that,

$$H(K|\Theta) \geq n\cdot(1-2\varepsilon) - h(\varepsilon) \quad \text{with} \quad \Theta \in \{\theta_1, \ldots, \theta_L\}\,, \tag{102}$$

where $h(\varepsilon)$ denotes the binary entropy. The relation (102) will be the basis for the information locking schemes presented in Sec. VI.H.3.

4. BB84 and six state measurements

In contrast to what we discussed so far, for applications in cryptography we usually need uncertainty relations for measurements that can be implemented locally. For example, for an n-qubit state we are interested in uncertainty relations for the set of 2^n different measurements given by measuring each qubit independently in one of the two Pauli bases $\sigma_{\mathbb{X}}$ or $\sigma_{\mathbb{Z}}$. Following Bennett and Brassard (1984) these measurements are often called BB84 measurements. Using the Maassen-Uffink bound (31) for two measurements iteratively we immediately find,

$$H(K^n|\Theta^n) \geq n\cdot\frac{1}{2} \quad \text{with} \quad \Theta^n \in \{\theta_1, \ldots, \theta_{2^n}\}\,. \tag{103}$$

This relation is already tight since there exist states that achieve equality. However, for applications the relevant measure is often not the Shannon entropy but the the min-entropy. The one qubit relation (72) is easily extended to n qubits as

$$H_{\min}(K^n|\Theta^n) \geq -n \cdot \log\left(\frac{1}{2} + \frac{1}{2\sqrt{2}}\right) \approx n \cdot 0.22 \quad \text{with } \Theta^n \in \{\theta_1, \ldots, \theta_{2^n}\}\,. \tag{104}$$

Again there exist states that achieve equality. More generally Ng *et al.* (2012) find for n qubit BB84 measurements and the Rényi entropy of order $\alpha \in (1, 2]$,

$$H_\alpha(K^n|\Theta^n) \geq n \cdot \frac{\alpha - \log\left(1 + 2^{\alpha-1}\right)}{\alpha - 1} \quad \text{with } \Theta^n \in \{\theta_1, \ldots, \theta_{2^n}\}\,, \tag{105}$$

where the conditioning is given as (see App. C),[9]

$$H_\alpha(K|\Theta) = \frac{\alpha}{1-\alpha} \log\left(\frac{1}{L}\sum_{j=1}^{L} 2^{\frac{1-\alpha}{\alpha} H_\alpha(K|\Theta=\theta_j)}\right)\,. \tag{106}$$

Similarly, we find for the set of 3^n different measurements given by measuring each qubit independently in one of the three Pauli bases $\sigma_{\mathbb{X}}$, $\sigma_{\mathbb{Y}}$, or $\sigma_{\mathbb{Z}}$ that

$$H(K^n|\Theta^n) \geq n \cdot \frac{2}{3} \quad \text{with } \Theta^n \in \{\theta_1, \ldots, \theta_{3^n}\}\,, \tag{107}$$

Following Bruß (1998) these measurements are often called six state measurements. The uncertainty relation (107) is the extension of (87) from one to n qubits. More general relations in terms of Rényi entropies were again derived by Ng *et al.* (2012).

Approximate extensions of all these relations when the measurements are not exactly given by the Pauli measurements $\{\sigma_{\mathbb{X}}, \sigma_{\mathbb{Y}}, \sigma_{\mathbb{Z}}\}$ are discussed in (Kaniewski *et al.*, 2014). We note that some extensions of the n qubit relations discussed above will be crucial for applications in two-party cryptography (Sec. VI.C).

5. General sets of measurements

Liu *et al.* (2015) give entropic uncertainty relations for general sets of measurements. Their bounds are qualitatively different than just combining (31) iteratively and sometimes become strictly stronger in dimension $d > 2$. For simplicity we only state the case of $L = 3$ measurements (in any dimension $d \geq 2$),

$$H(K|\Theta) \geq \frac{1}{3}\log\frac{1}{m} + \frac{2}{3}H(\rho_A) \quad \text{with } \Theta \in \{V^{(1)}, V^{(2)}, V^{(3)}\}\,, \tag{108}$$

and the multiple overlap constant

$$m := \max_k \left(\sum_j \max_i \left(c\left(v_i^1, v_j^2\right)\right) \cdot c\left(v_j^2, v_k^3\right)\right)\,, \tag{109}$$

and $\{|v_i^1\rangle\}$, $\{|v_j^2\rangle\}$, $\{|v_k^3\rangle\}$ are the eigenvectors of $V^{(1)}$, $V^{(2)}$, $V^{(3)}$, respectively.

Example 16. *For $d = 2$ and the full set of $d + 1 = 3$ MUBs given by the Pauli measurements this gives*

$$H(K|\Theta) \geq \frac{1}{3} + \frac{2}{3}H(A)_\rho \quad \text{with } \Theta \in \{\sigma_{\mathbb{X}}, \sigma_{\mathbb{Y}}, \sigma_{\mathbb{Z}}\}\,. \tag{110}$$

This bound is, however, weaker than (78) and (88). On the other hand, of course the whole point of the bound (108) is that in contrast to (78) and (88) it can be applied to any set of $L = 3$ measurements.

We refer to (Liu *et al.*, 2015) for a fully worked out example where their bound can become stronger than any bounds implied by two measurement relations.

6. Anti-commuting measurements

As already noted in Sec. III.D, many interesting measurements are not of the orthonormal basis form, but are more generally described by POVMs. One class of such measurements that generate maximally strong uncertainty relations are sets of anti-commuting POVMs with only two possible measurement outcomes. In more detail, we consider $\{\mathbb{X}_1, \ldots, \mathbb{X}_L\}$ of binary POVMs $\mathbb{X}_j = \{\mathbb{X}_j^0, \mathbb{X}_j^1\}$ that generate binary observables (Hermitian operators)

$$O_j := \mathbb{X}_j^0 - \mathbb{X}_j^1 \quad \text{with } [O_j, O_k]_+ = 2\delta_{jk}\,, \tag{111}$$

where, as in (59), $[\cdot,\cdot]_+$ denotes the anti-commutator. An example are the $L = 3$ qubit Pauli measurements $\{\sigma_{\mathbb{X}}, \sigma_{\mathbb{Y}}, \sigma_{\mathbb{Z}}\}$. In higher dimensions, the operators in (111) are the analogons of the Pauli operators. We are then interested in finding lower bounds on $H(K|\Theta)$ where

$$P_{K\Theta}(k, \mathbb{X}_j) := \frac{1}{L}\operatorname{tr}\left[\mathbb{X}_j^k \rho_A\right] \quad \text{with } k = 1, 2 \quad j = 1, \ldots, L\,. \tag{112}$$

[9] We emphasize that unlike in the unconditional case $H_2(K|\Theta) \neq H_{\text{coll}}(K|\Theta)$ and hence (82) is different from (106) for $\alpha = 2$.

For simplicity we only discuss the case of n qubit states for which we have sets of up to $2n+1$ many binary anti-commuting POVMs.[10] Wehner and Winter (2008) then show that

$$H(K|\Theta) \geq 1 - \frac{1}{L} \quad \text{with} \quad k = 1, 2$$
$$\text{for any subset } \Theta \subseteq \{\mathbb{X}_1, \ldots, \mathbb{X}_{2n+1}\} \text{ of size } L\,. \tag{113}$$

These relations are tight and reduce for the $L = 3$ qubit Pauli measurements $\{\sigma_{\mathbb{X}}, \sigma_{\mathbb{Y}}, \sigma_{\mathbb{Z}}\}$ to the bound (87). Similarly Wehner and Winter (2008) also find for the collision entropy,

$$\frac{1}{L} \sum_{\mathbb{X}_j \in \Theta} H_{\text{coll}}(K|\Theta = \mathbb{X}_j) \geq 1 - \log\left(1 + \frac{1}{L}\right)\,, \tag{114}$$

and the min-entropy,

$$\frac{1}{L} \sum_{\mathbb{X}_j \in \Theta} H_{\min}(K|\Theta = \mathbb{X}_j) \geq 1 - \log\left(1 + \frac{1}{\sqrt{L}}\right)\,. \tag{115}$$

These relations are again tight. Note however that the average over the basis choice is outside of the logarithm, whereas for the collision and the min-entropy the average is more naturally inside of the logarithm as, e.g., in (81) and in (104).

Example 17. *For the $L = 3$ qubit case (114) reduces to*

$$\frac{1}{3} \sum_{j=X,Y,Z} H_{\text{coll}}(K|\Theta = \sigma_j) \geq \log 3 - 1\,, \tag{116}$$

which, as seen by (84), is generally weaker than the corresponding bound implied by (81),

$$H_{\text{coll}}(K|\Theta) \geq \log 3 - 1 \quad \textit{with} \quad \Theta \in \{\sigma_{\mathbb{X}}, \sigma_{\mathbb{Y}}, \sigma_{\mathbb{Z}}\}\,. \tag{117}$$

Finally, we point to (Ver Steeg and Wehner, 2009) for the connection of the uncertainty relations in this section to Bell inequalities.

7. Mutually unbiased POVMs

In Sec. III.G.2 we presented that full sets of $d+1$ MUBs give rise to strong uncertainty relations, see, e.g., (80). However, for general dimension d we do not know if a full set of $d+1$ MUB always exists (see App. A for a discussion). Kalev and Gour (2014) offer the following generalization of MUBs to measurements that are not necessarily given by a basis. Two POVMs $\mathbb{X} = \{\mathbb{X}^x\}_{x=1}^d$ and $\mathbb{Z} = \{\mathbb{Z}^z\}_{z=1}^d$ on a d-dimensional quantum system are mutually unbiased measurements (MUMs) if for some $\kappa \in (1/d, 1]$,

$$\text{tr}\,[\mathbb{X}^x] = 1, \quad \text{tr}\,[\mathbb{Z}^z] = 1, \quad \text{tr}\,[\mathbb{X}^x \mathbb{Z}^z] = \frac{1}{d} \quad \forall x, z \tag{118}$$

$$\text{tr}\left[\mathbb{X}^x \mathbb{X}^{x'}\right] = \delta_{xx'} \cdot \kappa + (1 - \delta_{xx'}) \frac{1-\kappa}{d-1} \quad \forall x, x'$$
$$\text{and similarly for } z, z'\,. \tag{119}$$

In addition, a set of POVMs $\{\mathbb{X}_1, \ldots, \mathbb{X}_n\}$ of said form is called a set of MUMs if each POVM $\mathbb{X}_j$ is mutually unbiased to each other POVM $\mathbb{X}_k$, with $k \neq j$, in the set.

A straightforward example are again MUBs for which $\kappa = 1$.[11] The crucial observation of Kalev and Gour (2014) is that in any dimension d a full set of $d+1$ MUMs exists (see their paper for the explicit construction). Moreover, every full set of $d+1$ MUMs gives rise to a strong uncertainty relation,

$$H(K|\Theta) \geq \log(d+1) - \log(1+\kappa)$$
$$\text{with} \quad \Theta \in \{\mathbb{X}_1, \ldots, \mathbb{X}_{d+1}\}\,, \tag{120}$$

where the notation is as introduced in (112). This is in full analogy with (80) for a full set of $d+1$ MUBs. Tighter and state dependent versions of (120) as well as extensions to Rényi entropies can be found in (Chen and Fei, 2015; Rastegin, 2015b)

H. Fine-grained uncertainty relations

So far we have expressed uncertainty in terms of the von Neumann entropy and the Rényi entropy of the probability distribution induced by the measurement. Recall, however, that any restriction on the set of allowed probability distributions over measurement outcomes can be understood as an uncertainty relation, and hence there are many ways of formulating such restrictions. Thus, while in finite dimensions the Rényi entropies determine the underlying probability distribution of the measurement outcomes uniquely,[12] it is interesting to ask whether we can formulate more refined versions of uncertainty relations.

Suppose we perform L measurements labeled Θ on a preparation ρ, where each measurement has N outcomes. Fine-grained uncertainty relations (Oppenheim

[10] This is obtained by the unique Hermitian representation of the Clifford algebra via the Jordan-Wigner transformation (Dietz, 2006).

[11] The trivial example for which each POVM element is the maximally mixed state $\mathbb{1}/d$ is excluded because this would correspond to $\kappa = 1/d$.

[12] To see this, note that the cumulant generating function of the random variable $Z = -\log P_X(X)$ can be expressed in terms of the Rényi entropy of X, namely $g_Z(s) = H_{1+s}(X)$. The cumulants of Z and hence the distributions of Z and X are thus fully determined by the Rényi entropy in a neighborhood around $\alpha = 1$.

and Wehner, 2010) consist of a set of N^L equations which state that for all states we might measure

$$\sum_{\theta=1}^{L} P_\Theta(\Theta=\theta) P_X(X=x_\theta|\Theta=\theta) \leq \zeta_{x_1,\ldots,x_L}\,, \quad (121)$$

for all combinations of measurement outcomes $x_1,\ldots,x_L$ that are possible for the L different measurements, where $P_\Theta(\Theta=\theta)$ is the probability of choosing measurement labeled $\Theta=\theta$, and $0\leq\zeta_{x_1,\ldots,x_L}\leq 1$.

Note that whenever $\zeta_{x_1,\ldots,x_L}<1$, then we observe some amount of uncertainty, since it implies that we cannot simultaneously have $P_X(X=x_\theta|\Theta=\theta)=1$ for all θ. We remark that fine-grained uncertainty relations naturally give a bound on the min-entropy

$$H_{\min}(X|\Theta) = -\log \sum_{\theta=1}^{L} P_\Theta(\Theta=\theta) \max_{x_\theta} P_X(X=x_\theta|\Theta=\theta) \quad (122)$$

$$\geq -\log \max_{x_1,\ldots,x_L} \zeta_{x_1,\ldots,x_L}\,. \quad (123)$$

However, we note that fine-grained uncertainty relations are strictly more informative, and are closely connected to Bell non-locality (Oppenheim and Wehner, 2010). While not the topic of this survey, we note that a number of extensions of these fine-grained uncertainty relations are known (Dey *et al.*, 2013; Rastegin, 2015a; Ren and Fan, 2014).

I. Majorization approach to entropic uncertainty

Another way to capture uncertainty relations that relates directly to entropic ones is given by the majorization approach. Instead of sums of probabilities, we here look at products. The idea to derive entropic uncertainty relations via a majorization relation was pioneered by Partovi (2011) and later extended and clarified independently in (Friedland *et al.*, 2013) and (Puchała *et al.*, 2013). Let us recall the distributions P_X and P_Z resulting from the measurements $\mathbb{X}$ and $\mathbb{Z}$, respectively, of the state ρ_A as in (48). We denote by $P_X^\downarrow$ and $P_Z^\downarrow$ the corresponding reordered vectors such that the probabilities are ordered form largest to smallest.

1. Majorization approach

The main objective of this section is to find a vector that majorizes the Cartesian product of the two probability vectors $P_X^\downarrow$ and $P_Z^\downarrow$, namely we are looking for a probability distribution $\nu=\{\nu(1),\nu(2),\ldots,\nu(|X||Z|)\}$ such that the relation[13]

$$P_X^\downarrow \times P_Z^\downarrow \prec \nu \qquad \text{holds for all } \rho\in\mathcal{S}(\mathcal{H})\,. \quad (124)$$

Such a relation gives a bound on how spread out the product distribution $P_X^\downarrow \times P_Z^\downarrow$ must be. A simple and instructive example of a probability distribution ν satisfying the above relation can be constructed as follows. Consider the largest probability in the product distribution in (124), given by

$$p_1 := P_X^\downarrow(1)\cdot P_Z^\downarrow(1) = p_{\text{guess}}(X)\cdot p_{\text{guess}}(Z)\,. \quad (125)$$

We know that p_1 is always bounded away from 1 if the two measurements are incompatible, since it cannot be that both measurements have a deterministic outcome. For example, recall that we have (68) from Deutsch (1983), which gives

$$p_1 = p_{\text{guess}}(X)\cdot p_{\text{guess}}(Z) \leq b^2 =: \nu_1\,, \quad (126)$$

where b was defined in (40). As such, it is immediately clear that the vector $\nu^1=\{\nu_1,1-\nu_1,0,\ldots,0\}$ satisfies (124) and in fact constitutes a simple but non-trivial uncertainty relation.

Going beyond this observation, the works of (Friedland *et al.*, 2013) and (Puchała *et al.*, 2013) both present an explicit method to construct a sequence of vectors $\{\nu^k\}_{k=1}^{|X|-1}$ of the form

$$\nu^k = \{\nu_1,\nu_2-\nu_1,\ldots,1-\nu_{k-1},0,\ldots,0\}\,, \quad (127)$$

with $\nu^k \prec \nu^{k-1}$, that satisfy (124) and lead to tighter and tighter uncertainty relations. The expressions for ν_k are given in terms of an optimization problem and become gradually more difficult as k increases. We refer the reader to these papers for details on the construction.

2. From majorization to entropy

Entropic uncertainty relations for Rényi entropy follow directly from the above relation due to the fact that the Rényi entropy is Schur concave and additive. This implies that

$$P_X^\downarrow \times P_Z^\downarrow \prec \nu \implies H_\alpha(X)+H_\alpha(Z) \geq H_\alpha(V)\,, \quad (128)$$

where V is a random variable distributed according to the law ν. These uncertainty relations have a different flavor than the Maassen-Uffink relations in (35) since they provide a bound on the sum of the Rényi entropy of the same

[13] Recall the definition of majorization in Sec. III.A.3.

parameter. As a particular special case for $\alpha = \infty$, we get back Deutsch's uncertainty relation (Deutsch, 1983),

$$H(X) + H(Z) \geq H_{\min}(X) + H_{\min}(Z) \tag{129}$$
$$\geq -2\log b =: q_{\mathrm{D}}\,, \tag{130}$$

where the first inequality follows by the monotonicity of the Rényi entropy in the parameter α. However, an immediate improvement on this relation can be obtained by applying (128) directly for $\alpha = 1$, which yields

$$H(X) + H(Z) \geq h_{\mathrm{bin}}(b^2) =: q_{\mathrm{maj}}\,. \tag{131}$$

3. Measurements in random bases

An interesting special case for which a majorization based approach gives tighter bounds is for measurements in two bases $\mathbb{X}$ and $\mathbb{Z}$ related by a random unitary. Intuitively, we would expect such bases to be complementary. More precisely, for any measurement in a fixed basis $\mathbb{X}$ and $\mathbb{Z}$ related by a unitary drawn from the Haar measure on the unitary group, Adamczak *et al.* (2014) show that for the Masseen-Uffink bound (31) we have with probability going to one for $d \to \infty$,

$$H(X) + H(Z) \geq \log d - \log\log d\,. \tag{132}$$

However, they also show that a majorization based approach yields the tighter estimate

$$H(X) + H(Z) \geq \log d - C_1\,, \tag{133}$$

where $C_1 > 0$ is some constant. This is close to optimal since we have that with probability going to one for $d \to \infty$ (Adamczak *et al.*, 2014),

$$\log d - C_0 \geq H(X) + H(Z)\,, \tag{134}$$

for some constant $C_0 > 0$. It is an open question to determine the exact asymptotic behavior, i.e., the constant $C \in (C_0, C_1)$ that gives a lower and an upper bound.

4. Extensions

The majorization approach has also been extended to cover general POVMs and more than two measurements in (Friedland *et al.*, 2013). Moreover, a recent paper (Rudnicki *et al.*, 2014) discusses a related method, based on finding a vector that majorizes the ordered distribution $(P_X \cup P_Z)^{\downarrow}$, where $P_X \cup P_Z$ is simply the concatenation of the two probability vectors. This yields a further improvement on (131). Finally, an extension to uncertainty measures that are not necessarily Schur concave but only monotonic under doubly stochastic matrices was presented in (Narasimhachar *et al.*, 2015).

IV. UNCERTAINTY GIVEN A MEMORY SYSTEM

The uncertainty relations presented thus far are limited in the following sense: they do not allow the observer to have access to *side information*. Side information, also known as *memory*, might help the observer to better predict the outcomes of the $\mathbb{X}$ and $\mathbb{Z}$ measurements. It is therefore a fundamental question to ask: does the uncertainty principle still hold when the observer has access to a memory system? If so, what form does it take?

The uncertainty principle in the presence of memory is important for cryptographic applications and witnessing entanglement (Sec. VI). For example, in quantum key distribution, an eavesdropper may gather some information, store it in her memory, and then later use that memory to try to guess the secret key. It is crucial to understand whether the eavesdropper's memory allows her to break a protocol's security, or whether security is maintained. This is where general uncertainty relations that allow for memory are needed.

Furthermore, such uncertainty relations are also important for basic physics. For example, the quantum-to-classical transition is an area of physics where one tries to understand why and how quantum interference effects disappear on the macroscopic scale. This is often attributed to decoherence, where information about the system of interest S flows out to an environment E (Zurek, 2003). In decoherence, it is important to quantify the tradeoff between the flow of one kind of information, say $\mathbb{Z}$, to the environment versus the preservation of another kind of information, say $\mathbb{X}$, within the system S. Here one associates $\mathbb{X}$ with the "phase" information that is responsible for quantum interference. Hence, one can see how this ties back into the quantum-to-classical transition, since loss of $\mathbb{X}$ information would destroy the quantum interference pattern. In this discussion, system E plays the role of the memory, and hence uncertainty relations that allow for memory are essentially uncertainty relations that allow the system to interact with an environment. We will discuss this more in Sec. VI.F, in the context of interferometry experiments.

A. Classical versus quantum memory

With this motivation in mind, we now consider two different types of memories. First, we discuss the notion of a *classical memory*, i.e., a system B that has no more than classical correlations with the system A that is to be measured.

Example 18. *Consider a spin-1/2 particle A and a (macroscopic) coin B as depicted in Fig. 7(a). Suppose that we flip the coin to determine whether or not we prepare A in the spin-up state $|0\rangle$ or the spin-down state $|1\rangle$. Denoting the basis $\mathbb{Z} = \{|0\rangle, |1\rangle\}$ we see that B is perfectly*

correlated to this basis. That is, before the measurement of A the joint state is

$$\rho_{AB} = \frac{1}{2} \sum_{x\in\{0,1\}} |x\rangle\langle x|_A \otimes |x\rangle\langle x|_B \ . \tag{135}$$

Hence, if the observer has access to B then he can perfectly predict the outcome of the $\mathbb{Z}$ measurement on A. On the other hand, if we keep B hidden from the observer, then he can only guess the outcome of the $\mathbb{Z}$ measurement on A with probability $1/2$.

We conclude from Ex. 18 that indeed, having access to B reduces the uncertainty about $\mathbb{Z}$. However, notice that a classical memory B provides no help to the observer if he tries to guess the outcome of a measurement on A that is complementary to $\mathbb{Z}$. Consider now a more general memory, one that can have any kind of correlations with system A allowed by quantum mechanics. This is called a *quantum memory* or *quantum side information* (and includes classical memory as a special case). We remark that quantum memories are now becoming an experimental reality, e.g., see (Julsgaard *et al.*, 2004).

Example 18 (continued). *Consider two spin-1/2 particles A and B that are maximally entangled, say in the state*

$$|\psi\rangle_{AB} = \frac{1}{\sqrt{2}}\big(|00\rangle_{AB} + |11\rangle_{AB}\big) \ . \tag{136}$$

This is depicted in Fig. 7(b). Like for the classical memory in Ex. 18, giving the observer access to B allows him to perfectly predict the outcome of a $\mathbb{Z}$ measurement on A (by just measuring the $\mathbb{Z}$ observable on B). But in contrast to the case with classical memory, B can also be used to predict the outcome of a complementary measurement $\mathbb{X} = \{|+\rangle, |-\rangle\}$, with $|\pm\rangle = (|0\rangle \pm |1\rangle)/\sqrt{2}$, on A. This follows by rewriting the maximally entangled state (136) as

$$|\psi\rangle_{AB} = \frac{1}{\sqrt{2}}\big(|{+}{+}\rangle_{AB} + |{-}{-}\rangle_{AB}\big) \ , \tag{137}$$

which implies that the observer can simply measure the $\mathbb{X}$ basis on B to guess $\mathbb{X}$ on A.

The ideas described in Ex. 18 can be dated back to the famous EPR paper (Einstein *et al.*, 1935) and raise the question of whether we can still find nontrivial bounds on the uncertainty of complementary measurements when conditioning on quantum memory. In the rest of Sec. IV we analyze this interplay between uncertainty and quantum correlations quantitatively and present entropic uncertainty relations that allow the observer to have access to (quantum) memory. For that we first introduce measures of *conditional entropy*.

(a) Cartoon showing an electron spin whose $\mathbb{Z}$ component is correlated to a classical coin.

(b) Cartoon showing an electron spin whose $\mathbb{Z}$ and $\mathbb{X}$ components are respectively correlated to the $\mathbb{Z}$ and $\mathbb{X}$ components of another electron spin, i.e., a quantum memory.

FIG. 7 Comparison of classical and quantum memory.

B. Background: Conditional entropies

1. Classical-quantum states

Our main goal here is to describe the entropy of a measured (and thus classical) random variable from the perspective of an observer that has quantum memory. For this purpose, consider a classical register correlated with a quantum memory, modeled by a joint *classical quantum (cq) state*

$$\rho_{XB} = \sum_x P_X(x)\, |x\rangle\langle x|_X \otimes \rho_B^x \ . \tag{138}$$

Here, ρ_B^x is the quantum state of the memory system B conditioned on the event $X = x$. Formally, *quantum states* or *density operators* are positive semidefinite operators with unit trace acting on the Hilbert space B. In order to represent the joint system XB in the density operator formalism we also introduced an auxiliary Hilbert space X with fixed orthonormal basis $\{|x\rangle_X\}_x$.

2. Classical-quantum entropies

The interpretation of the min-entropy from (17) in terms of the optimal guessing probability gives a natural means to generalize the min-entropy to the setting with quantum memory. Clearly, an observer with access to the quantum memory B can measure out B to improve his guess. The optimal guessing probability for such an observer is then given by the optimization problem

$$p_{\text{guess}}(X|B) := \max_{\mathbb{X}_B} \sum_x P_X(x)\, \mathrm{tr}\left[\mathbb{X}_B^x \rho_B^x\right] , \quad \text{where } \mathbb{X}_B \text{ is a POVM on } B. \tag{139}$$

Consequently, the conditional min-entropy is defined as (König *et al.*, 2009; Renner, 2005),

$$H_{\min}(X|B) := -\log p_{\text{guess}}(X|B)\,. \tag{140}$$

This is our first measure of conditional entropy. It quantifies the uncertainty of the classical register X from the perspective of an observer with access to the quantum memory (or side information) B. The more difficult it is to guess the value of X, the smaller is the guessing probability and the higher is the conditional min-entropy.

The collision entropy from (16) can likewise be interpreted in terms of a guessing probability. Consider a generalization of the collision entropy to the case where the observer has a quantum memory B, which takes the form (Buhrman *et al.*, 2008),

$$H_{\text{coll}}(X|B) := -\log p^{\text{pg}}_{\text{guess}}(X|B)\,. \tag{141}$$

Here, the pretty good guessing probability is given by

$$p^{\text{pg}}_{\text{guess}}(X|B) := \sum_x P_X(x)\,\text{tr}\left[\Pi^x_B \rho^x_B\right], \quad \text{where } \Pi^x_B = P_X(x)\rho_B^{-1/2}\rho^x_B\rho_B^{-1/2}\,. \tag{142}$$

The Π^x_B are POVM elements corresponding to the so-called pretty good measurement. The name is due the fact that this measurement is close to optimal, in the sense that (Hausladen and Wootters, 1994),

$$p^2_{\text{guess}}(X|B) \le p^{\text{pg}}_{\text{guess}}(X|B) \le p_{\text{guess}}(X|B)\,. \tag{143}$$

That is, if the optimal guessing probability is close to one, then so is the pretty good guessing probability. Hence, $H_{\text{coll}}(X|B)$ quantifies how well Bob can guess X given that he performs the pretty good measurement on B. In particular this also implies that

$$H_{\min}(X|B) \le H_{\text{coll}}(X|B) \le 2H_{\min}(X|B)\,. \tag{144}$$

Finally, consider the Shannon entropy $H(X)$, whose quantum counterpart $H(\rho)$ is the von Neumann entropy as defined in (45). The von Neumann entropy of X conditioned on a quantum memory B, is defined as

$$H(X|B) := H(\rho_{XB}) - H(\rho_B)\,. \tag{145}$$

where ρ_{XB} is given by (138), and

$$\rho_B = \text{tr}_X\left[\rho_{XB}\right] = \sum_x P_X(x)\rho^x_B\,. \tag{146}$$

Although $H(X|B)$ does not have a direct interpretation as a guessing probability, it does have an operational meaning in information theory. For example, if Alice samples from the distribution P_X and Bob possesses system B, then $H(X|B)$ is the minimal information that Alice must send to Bob in order for Bob to determine the value of X. (More precisely, $H(X|B)$ is the minimal rate in bits per copy that Alice must send to Bob, in the asymptotic limit of many copies of the state ρ_{XB} (Devetak and Winter, 2003).)

3. Quantum entropies

The classical-quantum conditional entropy is merely a special case of the quantum conditional entropy. It is useful to introduce the latter here, since the quantum conditional entropy will play an important role in the following.

In the simplest case, the von Neumann conditional entropy of an arbitrary bipartite state ρ_{AB}, whose marginal is $\rho_B = \text{tr}_A(\rho_{AB})$, takes the form

$$H(A|B) := H(\rho_{AB}) - H(\rho_B)\,. \tag{147}$$

We remark that, in general, fully quantum conditional entropy can be negative.[14] This is a signature of entanglement, and in fact, the quantity $-H(A|B)$ provides a lower bound on the distillable entanglement (Devetak and Winter, 2005). We will discuss this connection further around (335) below.

The fully quantum min-entropy also has a connection to entanglement. Namely, it can be written as,

$$H_{\min}(A|B) := -\log\left(d_A \cdot F(A|B)\right), \tag{148}$$

where

$$F(A|B) := \max_{\mathcal{E}:B\to A'} F\Big((\mathcal{I}\otimes\mathcal{E})(\rho_{AB}), |\phi_{AA'}\rangle\langle\phi_{AA'}|\Big)\,. \tag{149}$$

Here $|\phi_{AA'}\rangle$ is a maximally entangled state of dimension $|A|$, F denotes the fidelity, and the maximization is over all quantum channels $\mathcal{E}$ that map B to A'. One can think of $F(A|B)$ as the recoverable entanglement fidelity, and in that sense $-H_{\min}(A|B)$ quantifies how close the state is to a maximally entangled state.

The fully quantum collision entropy can also be related to a recoverable entanglement fidelity, in close analogy to the discussion above for the classical-quantum case. Namely, we have (Berta *et al.*, 2014a),

$$H_{\text{coll}}(A|B) := -\log\left(d_A \cdot F^{\text{pg}}(A|B)\right), \tag{150}$$

where

$$F^{\text{pg}}(A|B) := F\big((\mathcal{I}\otimes\mathcal{E}^{\text{pg}})(\rho_{AB}), |\phi_{AA'}\rangle\langle\phi_{AA'}|\big)\,. \tag{151}$$

Here $\mathcal{E}^{\text{pg}}$ is the pretty good recovery map, whose action on an operator O is given by

$$\mathcal{E}^{\text{pg}}(O) = \text{tr}_B\left[(\mathbb{1}\otimes\rho_B^{-1/2}O\rho_B^{-1/2})\rho_{AB}\right]^T, \tag{152}$$

[14] This should not concern us further here; a consistent interpretation of negative entropies is possible in the context of quantum information processing (Horodecki *et al.*, 2006) and also in thermodynamics (del Rio *et al.*, 2011).

where ^{T} is the transpose. In analogy to (143), the pretty good recovery map is close to optimal (Barnum and Knill, 2002),

$$F^2(A|B) \leq F^{\rm pg}(A|B) \leq F(A|B)\,. \tag{153}$$

As in the classical case, the above conditional entropies emerge as special cases of Rényi entropies (Müller-Lennert *et al.*, 2013). We discuss this connection in Appendix C.

4. Properties of entropy

Section III.A.3 discussed properties of entropies, which, as we will see, are special cases of conditional entropies. Here we mostly discuss properties of the conditional von Neumann entropy $H(A|B)$, and only note that similar properties also hold for other conditional entropies such as $H_{\min}(A|B)$ and $H_{\rm coll}(A|B)$ (or more general Rényi entropies).

Firstly, the conditional entropy reduces to the unconditional entropy for product states. That is, for bipartite states of the form $\rho_{AB} = \rho_A \otimes \rho_B$, we have $H(A|B) = H(A)$. Secondly, note that the entropy of classical-quantum is non-negative,

$$H(X|B) \geq 0 \quad \text{for } X \text{ a classical register.} \tag{154}$$

In contrast, as noted above, the fully quantum entropy $H(A|B)$ can be negative.

A fundamental property is the so-called data-processing inequality. It says that the uncertainty of A conditioned on some system B never goes down if one processes system B, i.e., acts on B with a quantum channel $\mathcal{E} : B \to B'$. That is (Lieb and Ruskai, 1973; Lindblad, 1975; Uhlmann, 1977),

$$H(A|B) \leq H(A|B')\,. \tag{155}$$

This includes the case where system $B = B_1B_2$ is bipartite and the processing corresponds to discarding a subsystem, say B_2. In this case the data-processing inequality takes the form $H(A|B) \leq H(A|B_1)$. This inequality is intuitive in the sense that having access to more information can never increase the uncertainty.

Another useful property of conditional entropies is related to the monogamy of entanglement. This corresponds to the idea that the more A is entangled with B the less A is entangled with a purifying system C. Suppose that C is a system that purifies ρ_{AB}, i.e., $\rho_{ABC} = |\psi\rangle\langle\psi|$. Then, we have

$$H(A|B) = -H(A|C)\,. \tag{156}$$

Typically one associates entanglement with a negative conditional entropy, and indeed as discussed above, minus the conditional entropy lower bounds the distillable entanglement. In this sense, the relation in (156) captures the intuition of monogamy of entanglement. It implies that if ρ_{AB} has a negative conditional entropy, then ρ_{AC} must have a positive conditional entropy. So there is a trade-off between the entanglement present in ρ_{AB} and in ρ_{AC}.

The relation in (156) is called the duality relation, as it relates an entropy to its *dual* entropy. As we have seen the von Neumann entropy is dual to itself but in general the duality relation involves two different entropies. For example, the min-entropy is dual to the max-entropy,

$$H_{\max}(A|B) := -H_{\min}(A|C)\,. \tag{157}$$

We take (157) as the definition of the max-entropy, although an explicit expression in terms of the marginal ρ_{AB} can be derived (König *et al.*, 2009). More generally, the duality relation for the Rényi entropy family is given in App. C.2.

C. Classical memory uncertainty relations

We now have all the measures at hand to discuss uncertainty relations that allow for a memory system. Naturally, we begin with the simplest case of a classical memory. It turns out that uncertainty relations that allow for classical memory are often easy to derive from the uncertainty relations without memory, particularly for the Shannon entropy (Hall, 1995). Consider the conditional Shannon entropy, which can be written as

$$H(X|Y) = H(XY) - H(Y) = \sum_y P_Y(y) H(X|Y=y)\,. \tag{158}$$

Now consider some generic Shannon entropy uncertainty relation for measurements $\mathbb{X}^n$ and quantum states ρ_A:

$$\sum_n H(X_n) \geq q \quad \text{where} \quad P_{X_n}(x) = \langle \mathbb{X}_n^x | \rho_A | \mathbb{X}_n^x \rangle \quad \text{and } q > 0 \text{ state-independent}\,. \tag{159}$$

The goal is to extend this to quantum classical states ρ_{AY} where the classical memory Y holds some information about the preparation of the quantum marginal

$$\rho_{AY} = \sum_y P_Y(y)\, \rho_A^y \otimes |y\rangle\langle y|_Y \quad \text{with distributions} \quad P_{X_nY}(x,y) = P_Y(y) \langle \mathbb{X}_n^x | \rho_A^y | \mathbb{X}_n^x \rangle\,. \tag{160}$$

However, assuming that the uncertainty relation (159) holds for all quantum states, it holds in particular for each conditional state ρ_A^y associated with $Y = y$ in the classical memory Y. Averaging over y gives

$$\sum_y P_Y(y) \sum_n H(X_n|Y=y) \geq \sum_y P_Y(y) q = q\,. \tag{161}$$

Hence, we find by (158) that

$$\sum_n H(X_n) \geq q \quad \Longrightarrow \quad \sum_n H(X_n|Y) \geq q\,. \tag{162}$$

That is, any Shannon entropy uncertainty relation of the form (159) implies a corresponding uncertainty relation in terms of the conditional Shannon entropy of the form (162). Note that the conditional version (162) even provides a stronger bound, since by the data-processing inequality (155) conditioning on side information can only reduce uncertainty.

Example 19. *Consider a bipartite state ρ_{AB}, where Alice will measure system A in one of two bases $\mathbb{X}$ or $\mathbb{Z}$ and Bob will measure system B in the basis $\mathbb{Y}$. Then, the Maassen-Uffink relation (31) already implies*

$$H(X|Y) + H(Z|Y) \geq \log \frac{1}{c}\,, \tag{163}$$

for the distribution

$$P_{XY}(x,y) = \langle \mathbb{X}^x \otimes \mathbb{Y}^y | \rho_{AB} | \mathbb{X}^x \otimes \mathbb{Y}^y \rangle\,, \tag{164}$$

and analogously $P_{ZY}(z,y)$.

It is also worth noting that the classical memory Y can be considered multipartite, say, of the form $Y = Y_1 Y_2 ... Y_n$ (Cerf *et al.*, 2002; Renes and Boileau, 2009). Since by the data-processing inequality (155) discarding subsystems of Y can never reduce the uncertainty, (162) already implies that

$$\sum_n H(X_n) \geq q \quad \Longrightarrow \quad \sum_n H(X_n|Y_n) \geq q\,. \tag{165}$$

Example 20. *Consider a tripartite state ρ_{ABC}, where Alice will measure system A in one of two bases $\mathbb{X}$ or $\mathbb{Z}$, Bob will measure system B in the basis $\mathbb{Y}_B$, and the third party Charlie will measure system C in the basis $\mathbb{Y}_C$. Then, the Maassen-Uffink relation (31) implies*

$$H(X|Y_B) + H(Z|Y_C) \geq \log \frac{1}{c}\,. \tag{166}$$

This relation is reminiscent of the scenario in quantum key distribution. Namely, if Alice and Bob verify that $H(X|P_B)$ is close to zero, then (166) implies that Charlie is fairly ignorant about Z. That is, $H(Z|Y_C)$ is roughly $\log(1/c)$ or larger. We emphasize, however, that (166) cannot be used to prove security against general quantum memory eavesdropping attacks.

D. Bipartite quantum memory uncertainty relations

1. Guessing game with quantum memory

Let us now make explicit what the guessing game (see Section III.F) looks like when we allow quantum memory. Specifically, the rules of the game now allow Bob to keep a *quantum* memory system in order to help him guess Alice's measurement outcome. This is illustrated in Fig. 8:

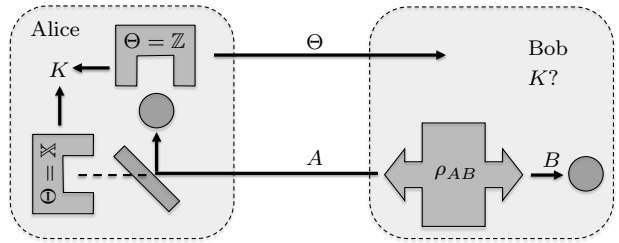


FIG. 8 Cartoon showing the guessing game in the presence of a memory system. First, a source prepares AB in state ρ_{AB}, and then sends system A to Alice and system B to Bob. Second, Alice performs either the $\mathbb{X}$ or $\mathbb{Z}$ measurement on A, and then announces the measurement choice Θ to Bob. Bob wins the guessing game if he correctly produces K. The question is thus: how uncertain is Bob about Alice's measurement outcome K, given that he has access to B and Θ.

1. Bob prepares a bipartite quantum system AB in a state ρ_{AB}. He sends system A to Alice while he keeps the system B.
2. Alice performs one of two possible measurements, $\mathbb{X}$ or $\mathbb{Z}$, on A, and stores the outcome in the classical register K. She communicates her choice to Bob.
3. Bob's task is to guess K.

Note that in this game, Bob can make an *educated* guess based on his quantum memory B.

Example 21. *Let the A system be one qubit and Alice's two measurements given by $\sigma_{\mathbb{X}}$ and $\sigma_{\mathbb{Z}}$. Then Bob can win the game with probability one by just preparing the maximally entangled state and using the strategy described in Ex. IV.A.*

This example illustrates the power of a quantum memory, and in particular, one that is *entangled* with the system being measured. At first sight, this might seem to violate the usual notion of the uncertainty principle. However, of course, it does not. What it illustrates is that the usual formulations of the uncertainty principle, such as the Kennard-Robertson relation (2) or Maassen-Uffink relation (31), are not about conditional uncertainty. The relations (2) and (31) are perfectly valid but limited in this sense.

2. Measuring in two orthonormal bases

Let us first discuss how the Maassen-Uffink relation (31) can be extended to the setup when the observer has a quantum memory. Note that Ex. 18 and 21 illustrate that the bound in the uncertainty relation must

become trivial in the case where Bob's memory is *maximally entangled* to Alice's system. On the other hand, we know that the bound must be non-trivial when Bob has no memory, since this corresponds to the situation covered by the Maassen-Uffink relation (31). Likewise if Bob has a memory that is only *classically* correlated to Alice's system, then we already saw in (163) that the Maassen-Uffink relation can be extended. Therefore, it becomes clear that we need a *state-dependent* extension: a bound that becomes weaker as Bob's memory is more entangled with Alice's system. Indeed, Berta *et al.* (2010) have proven the following uncertainty relation: for any bipartite state ρ_{AB} and any orthonormal bases $\mathbb{X}$ and $\mathbb{Z}$,

$$H(X|B) + H(Z|B) \geq \log \frac{1}{c} + H(A|B)\,, \tag{167}$$

with the overlap c as in (32). Here, the conditional entropy $H(X|B)$ is evaluated on the classical quantum state

$$\rho_{XB} = \sum_x |x\rangle\langle x|_X \otimes \big(\langle \mathbb{X}^x| \otimes \mathbb{1}_B\big)\rho_{AB}\big(|\mathbb{X}^x\rangle \otimes \mathbb{1}_B\big)\,, \tag{168}$$

and similarly for $H(Z|B)$. The classical quantum conditional entropies $H(X|B)$ and $H(Z|B)$ quantify Bob's uncertainty about X and Z respectively, given that Bob has access to the quantum memory B.

The quantity $H(A|B)$ on the right-hand side of (167) makes the bound state-dependent. We already mentioned around (147) that $-H(A|B)$ is a quantifier of the entanglement present in ρ_{AB}. For maximally entangled states we have $-H(A|B) = \log d_A$, whereas for all separable (i.e., non-entangled) states we have $H(A|B) \geq 0$.

Example 22. *Let us explore in more detail how the bound* (167) *behaves for some illustrative cases:*

1. *For maximally entangled states we get*

$$\log \frac{1}{c} + H(A|B) = \log \frac{1}{c} - \log d_A \leq 0\,, \tag{169}$$

and hence the bound becomes trivial. This is as expected from the guessing game example discussed in Section IV.D.1.

2. *For the case when Bob has no memory (i.e., B is trivial),* (167) *reduces to* (47),

$$H(X) + H(Z) \geq \log \frac{1}{c} + H(\rho_A)\,. \tag{170}$$

This is essentially the Maassen-Uffink relation but strengthened for mixed states.

3. *If B is not entangled with A (i.e., the state is separable), then $H(A|B) \geq 0$. Hence, we obtain*

$$H(X|B) + H(Z|B) \geq \log \frac{1}{c}\,. \tag{171}$$

This last example illustrates that (167) has applications for entanglement witnessing. More precisely, first note that by the data-processing inequality (155), (167) also implies

$$H(X|Y_B) + H(Z|W_B) \geq \log \frac{1}{c} + H(A|B) \quad \text{with } \mathbb{Y}_B \text{ and } \mathbb{W}_B \text{ measurements on } B. \tag{172}$$

Now violating this $\log(1/c)$ lower bound implies that the state ρ_{AB} must be entangled. We discuss this in detail in Sec. VI.D.

Using the following extension of the notation from Sec. III.F to quantum memory,

$$\rho_{K\Theta B} := \frac{1}{2}\sum_k \sum_{j=X,Z} |k\rangle\langle k|_K \otimes |j\rangle\langle j|_\Theta \otimes \Big(\langle k|U_j^\dagger \otimes 1_B\Big)\rho_{AB}\Big(U_j|k\rangle \otimes 1_B\Big)\,, \tag{173}$$

we can rewrite (167) as

$$H(K|B\Theta) \geq \frac{1}{2}\left(\log \frac{1}{c} + H(A|B)\right)\,. \tag{174}$$

This is the extension of (64) to quantum memory. Writing the relation in this way also makes a connection to the guessing game discussed in Sec. IV.D.1, see Fig. 8. We point to Sec. IV.D.7 for a partial extension of (174) in terms of the more operational min-entropy.

Let us take a step back and look at the history that led up to the uncertainty relation (167). Arguably the first work on uncertainty relations with quantum memory was by Christandl and Winter (2005). Their formulation was restricted to bases that are related by the Fourier transform but their work captures similar intuition as (167). The main difference, however, is that their relations are formulated for quantum channels rather than for quantum states. We discuss quantum channel uncertainty relations in Sec. IV.G.

Renes and Boileau (2009) gave the first quantum memory uncertainty relation in terms of the quantum state perspective. However, instead of bipartite states ρ_{AB}, they considered tripartite states ρ_{ABC}. We discuss entropic uncertainty relations for tripartite states in the next subsection (Sec. IV.E). Moreover, there is a close connection between tripartite and bipartite uncertainty relations. In fact, as we will also discuss in Sec. IV.E, Renes and Boileau (2009) conjectured a tripartite uncertainty relation that is equivalent to (167). Sec. IV.E also discusses the proof of quantum memory uncertainty relations such as (167), and notes that the tripartite formulation of (167) naturally generalizes to the Rényi entropy family.

3. Arbitrary measurements

Here we discuss some generalizations of (167) for arbitrary measurements. (See also Sec. IV.E.6 for a discussion of how the tripartite formulation generalizes to arbitrary measurements.)

Recall from Sec. III.D that the Maassen-Uffink relation generalizes to POVMs with the overlap c given by (49). In contrast, (167) does not hold for general POVMs with c as in (49). This can be remedied in two ways. The approach by Frank and Lieb (2013a) leads to a relation of the form (167) using a weaker complementarity factor. We have

$$H(X|B) + H(Z|B) \geq \log \frac{1}{c''} + H(A|B) \tag{175}$$

$$\text{where} \quad c'' = \max_{x,z} \operatorname{tr}\left[\mathbb{X}^x \mathbb{Z}^z\right]. \tag{176}$$

Note that $c'' \geq c$ in general and that c'' reduces to c for measurements in bases. However, one may argue that the form (175)–(176) is not natural if we consider general projective measurements or POVMs. This is best explained by means of an example (Furrer *et al.*, 2014).

Example 23. *Consider a quantum system A comprised of two qubits, A_1 and A_2, where A_1 is maximally entangled with a second qubit, B, and A_2 is in a fully mixed state, in product with A_1 and B. We employ rank-two projective measurements $\mathbb{X}_{A_1}$ and $\mathbb{Z}_{A_1}$ which measure A_1 in two mutually unbiased bases and leave A_2 intact. Analogously, we employ $\mathbb{X}_{A_2}$ and $\mathbb{Z}_{A_2}$ which measure A_2 in two mutually unbiased bases and leave A_1 intact. Evaluating the terms of interest for the measurement pairs $\{\mathbb{X}_{A_1}, \mathbb{Z}_{A_1}\}$ and $\{\mathbb{X}_{A_2}, \mathbb{Z}_{A_2}\}$ yields $c = \frac{1}{2}$ and $c'' = 1$ in both cases. Moreover, we find that*

$$H(A|B) = H(A_1|B) + H(A_2) = -1 + 1 = 0\,. \tag{177}$$

Hence, the right hand side of the Frank and Lieb relation (175) vanishes for both measurement pairs. Indeed, if the maximally entangled system A_1 is measured, we find that

$$H(X|B) + H(Y|B) = 0\,, \tag{178}$$

and the bound in (175) becomes an equality for the measurement pair $\{\mathbb{X}_{A_1}, \mathbb{Z}_{A_1}\}$. On the other hand, if A_2 is measured instead, we find that

$$H(X|B) + H(Y|B) = 2\,, \tag{179}$$

and the bound is far from tight for the measurement pair $\{\mathbb{X}_{A_2}, \mathbb{Z}_{A_2}\}$.

Examining this example, it is clear that the expected uncertainty depends strongly on which of the two systems is measured. Or, more generally, on how much entanglement is consumed in the measurement process. However, this information is not taken into account by the overlaps c or c'', nor by the entanglement of the initial state as measured by $H(A|B)$. Example 23 suggests that (167) can be generalized by considering the difference in entanglement of the state before and after measurement. In fact, Tomamichel (2012) shows the bipartite uncertainty relation

$$H(X|B) + H(Z|B) \geq \log \frac{1}{c'} + H(A|B) - \min\Big\{H(A'|XB),\, H(A'|ZB)\Big\}\,, \tag{180}$$

with c' given by (51). The entropies $H(A'|XB)$ and $H(A'|ZB)$ are evaluated for the post-measurement states

$$\rho_{XA'B} = \sum_x |x\rangle\langle x|_X \otimes (\mathbb{X}_A^x \otimes \mathbb{1}_B)\rho_{AB}(\mathbb{X}_A^x \otimes \mathbb{1}_B) \tag{181}$$

$$\rho_{ZA'B} = \sum_z |z\rangle\langle z|_Z \otimes (\mathbb{Z}_A^z \otimes \mathbb{1}_B)\rho_{AB}(\mathbb{Z}_A^z \otimes \mathbb{1}_B)\,, \tag{182}$$

respectively. (We use $A' = A$ to denote the system A after measurement to avoid confusion.) Notably, the terms $H(A'|XB)$ vanish for measurement given by a basis since in this case the state of A' is pure conditioned on X.

Example 23 (continued). *It is straightforward to see that if A_1 (A_2) is measured, the average entanglement left in the post-measurement state measured by the von Neumann entropy is given by $H(A_2|B)$ ($H(A_1|B)$). Hence, (180) turns into*

$$H(X|B) + H(Y|B) \geq \log \frac{1}{c} + \big(H(A|B) - H(A'|B)\big)\,, \tag{183}$$

where A' corresponds to A_2 (A_1). This inequality is tight for both measurements.

4. Multiple measurements

The basic goal here is to lift some of the relations in Sec. III.G to quantum memory. A general approach for deriving such relations is provided in (Dupuis *et al.*, 2015). The extensions will be to the bipartite setting as in (167). As in the unconditional case (cf. Sec. III.G.1), relations for two measurements already imply bounds for larger sets of measurements. For example for two qubit states ρ_{AB} and the Pauli measurements performed on the A-system we find by the simple iterative application of the bound (167) for the measurement pairs $\{\sigma_{\mathbb{X}}, \sigma_{\mathbb{Y}}\}$, $\{\sigma_{\mathbb{X}}, \sigma_{\mathbb{Z}}\}$, and $\{\sigma_{\mathbb{Y}}, \sigma_{\mathbb{Z}}\}$ that

$$H(K|B\Theta) \geq \frac{1}{2} + \frac{1}{2}H(A|B) \;\text{ with }\; \Theta \in \{\sigma_{\mathbb{X}}, \sigma_{\mathbb{Y}}, \sigma_{\mathbb{Z}}\}\,. \tag{184}$$

Here we use the following extension of the notation from Sec. III.G to quantum memory,

$$\rho_{K\Theta B} := \frac{1}{3}\sum_{k=1,2}\sum_{j=X,Y,Z} |k\rangle\langle k|_K \otimes |j\rangle\langle j|_\Theta \otimes \Big(\langle k|U_j^\dagger \otimes 1_B\Big)\rho_{AB}\Big(U_j|k\rangle \otimes 1_B\Big) . \quad (185)$$

Note that alternatively the left-hand side of (184) might also be written as

$$H(K|B\Theta) = \frac{1}{3}\Big(H(K|B\Theta=\sigma_{\mathbb{X}}) + H(K|B\Theta=\sigma_{\mathbb{Y}}) + H(K|B\Theta=\sigma_{\mathbb{Z}})\Big) . \quad (186)$$

where

$$\rho_{KB|\Theta=\sigma_{\mathbb{X}}} := \sum_{k=1,2} |k\rangle\langle k|_K \otimes \Big(\langle k|U_X^\dagger \otimes 1_B\Big)\rho_{AB}\Big(U_X|k\rangle \otimes 1_B\Big) , \quad (187)$$

and similarly for $\sigma_{\mathbb{Y}}, \sigma_{\mathbb{Z}}$. We refer to Sec. III.A for an in depth discussion of conditional entropies. The goal in all of Sec. IV.D.4 will be to find uncertainty relations that are stronger than any bounds that can be derived directly from relations for two measurements.

5. Complex projective two-designs

Berta *et al.* (2014a) show that the uncertainty equality (81) in terms of the collision entropy for a full set of MUBs also holds with quantum memory. That is, for any bipartite state ρ_{AB} with a full set of $d+1$ MUBs on the d dimensional A-system,

$$H_{\rm coll}(K|B\Theta) = \log(d+1) - \log\Big(2^{-H_{\rm coll}(A|B)}+1\Big) \quad \text{with } \Theta \in \{\theta_1,\ldots,\theta_{d+1}\} . \quad (188)$$

Here, as in (173), we use the notation,

$$\rho_{K\Theta B} := \frac{1}{d+1}\sum_{k=1}^{d}\sum_{j=1}^{d+1} |k\rangle\langle k|_K \otimes |j\rangle\langle j|_\Theta \otimes \Big(\langle k|U_j^\dagger \otimes 1_B\Big)\rho_{AB}\left(U_j|k\rangle \otimes 1_B\right) . \quad (189)$$

Example 24. *For the qubit Pauli measurements* (188) *yields:*

$$H_{\rm coll}(K|B\Theta) = \log 3 - \log\Big(2^{-H_{\rm coll}(A|B)}+1\Big) \quad \textit{with } \Theta \in \{\sigma_{\mathbb{X}},\sigma_{\mathbb{Y}},\sigma_{\mathbb{Z}}\} . \quad (190)$$

Since the collision entropy has an interpretation in terms of the pretty good guessing probability (141),

$$H_{\rm coll}(X|B) = -\log p^{\rm pg}_{\rm guess}(X|B) , \quad (191)$$

and the pretty good recovery map (150),

$$H_{\rm coll}(A|B) = -\log\big(d_A \cdot F^{\rm pg}(A|B)\big) , \quad (192)$$

the uncertainty equality (188) can be understood as an entanglement-assisted game of guessing complementary measurement outcomes (as described in Sec. IV.D.1). Namely, we can rewrite (188) as,

$$p^{\rm pg}_{\rm guess}(K|B\Theta) = \frac{d\cdot F^{\rm pg}(A|B)+1}{d+1} . \quad (193)$$

This gives a one-to-one relation between uncertainty (certainty) as measured by $p^{\rm pg}_{\rm guess}(K|B\Theta)$ and the absence (presence) of entanglement as measured by $F^{\rm pg}(A|B)$. In contrast, quantum memory assisted uncertainty relations for two measurements as, e.g., in (174) only provide us with a connection between uncertainty and entanglement in one direction. Namely, that low uncertainty implies the presence of entanglement (cf. Sec. VI.D).

The uncertainty equality (188) is derived by extending the proof from (Ballester and Wehner, 2007) that made use of the fact that a full set of mutually unbiased bases generates a complex projective two-design (Klappenecker and Rotteler, 2005). From this, it is also immediate that an equality as (188) holds for other complex projective two-designs as well. This includes in particular so-called symmetric informationally complete positive operator valued measures: SIC-POVMs.[15] More precisely, any SIC-POVM

$$\left\{\frac{1}{d}|\psi_k\rangle\langle\psi_k|\right\}_{k=1}^{d^2} \quad (194)$$

gives rise to the uncertainty equality

$$H_{\rm coll}(K|B\Theta) = \log\Big(d(d+1)\Big) - \log\Big(2^{-H_{\rm coll}(A|B)}+1\Big) \quad \text{with } \Theta \in \{\theta_1,\ldots,\theta_{d+1}\} . \quad (195)$$

Other examples that generate complex projective two-designs are unitary two-designs.[16] This includes in particular the Clifford group for n qubit systems.

Berta *et al.* (2014b) also show that the relation (188) for a full set of $d+1$ MUBs generates the following relation in terms of the von Neumann entropy,

$$H(K|B\Theta) \geq \log(d+1) - 1 + \min\Big\{0, H(A|B)\Big\} \quad \text{with } \Theta \in \{\theta_1,\ldots,\theta_{d+1}\} . \quad (196)$$

[15] We refer to (Renes *et al.*, 2004) for more on these objects.
[16] We refer to (Dankert *et al.*, 2009) for more on these objects.

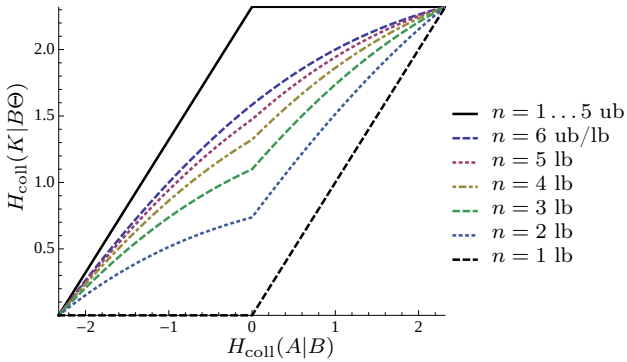


FIG. 9 For $d = 5$ and a various number of MUBs $n \leq d{+}1 = 6$ we plot the lower bounds (lb) on the entropic uncertainty $H_{\rm coll}(K|\Theta)$ from (197) as a function of $H_{\rm coll}(A|B)$. Moreover, for $n < 6$ we only have the trivial upper bound (ub) on $H_{\rm coll}(K|B\Theta)$, whereas for $n = 6$ the lower and upper coincide as in (188).

This corresponds to the generalization of (80) to quantum memory. Note that the entropy dependent term on the right-hand side only makes a contribution if the conditional entropy $H(A|B)$ is negative. This is consistent with (80).

For smaller sets of $L < d{+}1$ MUBs, Berta *et al.* (2014a) extend (93) to quantum memory,

$$H_{\rm coll}(K|B\Theta) \geq \begin{cases} -\log \frac{d\cdot 2^{-H_{\rm coll}(A|B)}+L-1}{L\cdot d} & \text{for } H_{\rm coll}(A|B) \geq 0 \\ -\log \frac{d+(L-1)2^{-H_{\rm coll}(A|B)}}{L\cdot d} & \text{for } H_{\rm coll}(A|B) < 0 \end{cases} \quad \text{with } \Theta \in \{\theta_1, \ldots, \theta_L\}\,. \tag{197}$$

Moreover, for all d and L there exist states that achieve equality. Note that for $L = d + 1$ the distinction of cases in (197) collapses and furthermore also become an upper bound as shown in (188). In Fig. 9 we illustrate this by means of an example for $d = 5$ (with $L \leq 6$).

6. Measurements in random bases

In the unconditional case we found that measurements in random bases lead to strong uncertainty relations as, e.g., in (100). Hence, we might expect that we can generalize this to quantum memory,

$$H(K|B\Theta) \overset{?}{\geq} O\left(\log d \cdot \left(1 - \frac{1}{L}\right)\right) + \min\left\{0, H(A|B)\right\} \quad \text{with } \Theta \in \{\theta_1, \ldots, \theta_L\} \text{ chosen at random}\,. \tag{198}$$

Unfortunately, the previous works (Fawzi *et al.*, 2011) and (Adamczak *et al.*, 2014) make use of measure concentration and ε-nets arguments that seem to fail for quantum memory. It is, however, possible to use some of the techniques from (Berta *et al.*, 2014a) based on operator Chernoff bounds to derive relations of the form (198). The downside is that we only get strong uncertainty relations for a large number L of measurements, that is,

$$L \geq O\Big(d \log(d)\Big)\,. \tag{199}$$

We conclude that it is an open problem to show the existence of small(er) sets of $L > 2$ measurements that generate strong uncertainty relations that hold with quantum memory.

7. BB84 and six state measurements

Let us now consider uncertainty relations as they are of relevance in quantum cryptography. For historical reasons we start with the n qubit six state measurements and only discuss the BB84 measurements later. For the six state measurements, Berta (2013) shows that for any bipartite state ρ_{A^nB} with the A^n-system given by n qubits,

$$H_{\rm coll}(K^n|B\Theta^n) \geq n \cdot \log 3 - \log\left(2^{-H_{\rm coll}(A^n|B)} + 1\right) \quad \text{with } \Theta^n \in \{\theta_1, \ldots, \theta_{3^n}\}\,. \tag{200}$$

This extends (190) from one to n qubits. The bound (200) also implies a similar relation in terms of the von Neumann entropy (Berta *et al.*, 2014b), extending (107) to

$$H(K^n|B\Theta^n) \geq n \cdot \log 3 + \min\left\{0, H(A^n|B)_\rho\right\} \quad \text{with } \Theta^n \in \{\theta_1, \ldots, \theta_{3^n}\}\,. \tag{201}$$

More importantly, Dupuis *et al.* (2015) improved (200) to the conceptually different bound

$$H_{\rm coll}(K^n|B\Theta^n) \geq n \cdot \gamma_{6s}\left(\frac{H_{\rm coll}(A^n|B)}{n}\right) - 1\,, \tag{202}$$

where

$$\gamma_{6s}(x) := \begin{cases} x & \text{if } x \geq \frac{\log 3}{2} \\ f^{-1}(x)\log 3 & \text{if } x < \frac{\log 3}{2} \end{cases} \quad \text{with} \quad f(x) = h(x) + x\log 3 - 1\,, \tag{203}$$

and $h(x)$ denotes the binary entropy function. Using the equivalence between the collision and the min-entropy from (144) this readily implies a relation as (202), but with both the collision entropy terms $H_{\rm coll}$ replaced with min-entropy terms $H_{\min}$. Importantly, this variant remains non-trivial for all values of $H_{\min}(A^n|B)$. What's more, (Dupuis *et al.*, 2015) establishes a powerful meta theorem that can be used to derive uncertainty relations also for other kinds of measurements, which find applications in quantum cryptography (Ribeiro and Grosshans, 2015).

For the n qubit BB84 measurements Dupuis *et al.* (2015) find

$$H_{\rm coll}(K^n|B\Theta^n) \geq n \cdot \gamma_{BB84}\left(\frac{H_{\rm coll}(A^n|B)}{n}\right) - 1 \quad \text{with} \quad \Theta^n \in \{\theta_1, \ldots, \theta_{2^n}\}\,. \tag{204}$$

where

$$\gamma_{BB84}(x) := \begin{cases} x & \text{if } x \geq \frac{1}{2} \\ g^{-1}(x) & \text{if } x < \frac{1}{2} \end{cases} \quad \text{with} \quad g(x) = h(x) + x - 1\,. \tag{205}$$

Again using the equivalence between the collision and the min-entropy from (144), we get a relation as (204) but with both the collision entropy terms $H_{\rm coll}$ replaced with min-entropy terms $H_{\min}$. We note that this is even new for one qubit ($n = 1$) and only the two measurements $\Theta \in \{\sigma_{\mathbb{X}}, \sigma_{\mathbb{Z}}\}$. The relation (204) and its min-entropy analogue can be understood in terms of the bipartite guessing game with quantum memory as mentioned in Sec. IV.D.1.

8. General sets of measurements

Liu *et al.* (2015) also give bipartite entropic uncertainty relations with quantum memory for general sets of measurements. Again for simplicity we only state the case of $L = 3$ observables (in any dimension $d \geq 2$). We find as the direct extension of (108),

$$H(K|B\Theta) \geq \frac{1}{3}\log\frac{1}{m} + \frac{2}{3}H(A|B) \quad \text{with} \quad \Theta \in \{V^{(1)}, V^{(2)}, V^{(3)}\}\,, \tag{206}$$

where the multiple overlap constant m is defined as in (109). Like in the unconditional case, this has to be compared with the bounds implied by two measurement relations as in (184). We refer to (Liu *et al.*, 2015) for a fully worked out example where (206) can become stronger than any bounds implied by two measurement relations.

E. Tripartite quantum memory uncertainty relations

1. Uncertainty relation with quantum memory

The physical scenario corresponding to tripartite uncertainty relations is shown in Fig. 10. Suppose there is a source that outputs ABC in state ρ_{ABC}. Systems A, B, and C are respectively sent to Alice, Bob, and Charlie. Then Alice performs either the $\mathbb{X}$ or $\mathbb{Z}$ measurement. If she measures $\mathbb{X}$, then Bob's goal is to minimize his uncertainty about X. If she measure $\mathbb{Z}$, then Charlie's goal is to minimize his uncertainty about Z. Renes and Boileau (2009) consider exactly this scenario but restrict to the case where the $\mathbb{X}$ and $\mathbb{Z}$ bases are related by the Fourier matrix F,

$$|\mathbb{X}^x\rangle = F|\mathbb{Z}^x\rangle \quad \text{with} \quad F = \sum_{z,z'} \frac{\omega^{-zz'}}{\sqrt{d}} |\mathbb{Z}^z\rangle\langle\mathbb{Z}^{z'}| \quad \text{where} \quad \omega = e^{2\pi i/d}\,. \tag{207}$$

Notice that this makes $\mathbb{X}$ and $\mathbb{Z}$ mutually unbiased, although in general not all pairs of MUBs are related by the Fourier matrix. They quantified Bob's and Charlie's uncertainty in terms of the conditional entropies $H(X|B)$ and $H(Z|C)$ respectively, and proved that any tripartite state ρ_{ABC} satisfies the relation

$$H(X|B) + H(Z|C) \geq \log d\,. \tag{208}$$

Here d is the dimension of the A system and the classical-quantum states ρ_{XB} and ρ_{ZC} are defined similarly as in (168). Renes and Boileau (2009) also conjectured that this relation generalizes to arbitrary measurements given by bases,

$$H(X|B) + H(Z|C) \geq \log\frac{1}{c}\,, \tag{209}$$

with the overlap c as in (32). Intuitively, what (208) says is that the more Bob knows about Z, the less Charlie knows about X, and vice-versa. This is a signature of the well-known trade-off *monogamy of entanglement*, which roughly says that the more Bob is entangled with Alice, the less he is with Charlie.[17] The trade-off described by (208) and (209) can be viewed as a more fine-grained notion of this monogamy, namely the monogamy appears at the level of measurement pairs $(\mathbb{X}, \mathbb{Z})$.

Also note that (209) implies both the Maassen-Uffink relation (31) and its classical memory extension (166), due to the data-processing inequality (155). That is,

$$H(X|B) \leq H(X|Y) \leq H(X)\,, \tag{210}$$

for any measurement $\mathbb{Y}$ on B. As we will see in Sec. IV.E.3 the quantum memory extension (209) is strictly stronger than the classical memory extension (166).

2. Proof of quantum memory uncertainty relations

The quantum memory uncertainty relation (209) was first proved by Berta *et al.* (2010). Although these authors stated their relation in the bipartite form (167), the two relations are equivalent.

---

[17] We refer to (Horodecki *et al.*, 2009) for an in-depth review about entanglement.

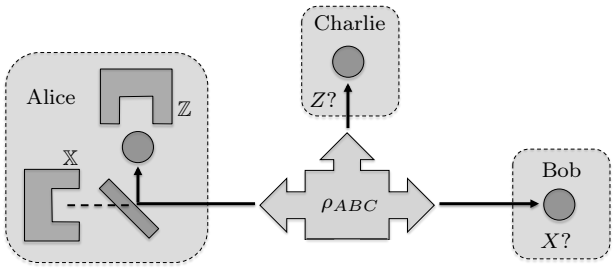


FIG. 10 Cartoon showing the tripartite memory scenario. First, a source prepares ABC in state ρ_{ABC}, then sends A to Alice, B to Bob, and C to Charlie. Then, Alice measures either $\mathbb{X}$ or $\mathbb{Z}$ on A and asks: how uncertain is Bob about her X outcome, given B, and how uncertain is Charlie about her Z outcome, given C? As shown in (209) there is a trade-off that is quantified by the complementarity of the measurements X and Z. We can also interpret this scenario as a guessing game, also called a monogamy game, that may look slightly different at first glance: In this game, Bob and Charlie play against Alice. They prepare ρ_{ABC} where they send A to Alice, Bob keeps B and Charlie keeps C. Alice then randomly chooses a measurement obtaining outcome K. Afterwards, she sends her choice of basis to Bob and Charlie. They win if and only if Bob outputs K, and at the same time Charlie outputs K. An uncertainty relation says that Bob and Charlie cannot win perfectly, that is, not both Bob and Charlie can produce K. This game measures the same kind of uncertainty, explicitly exploiting the monogamy of entanglement: if Bob produces $K = X$ correctly if Alice measured $\mathbb{X}$, then this is a certificate that Charlie cannot produce a good guess of $K = Z$ if Alice measured $\mathbb{Z}$.

The equivalence between the bipartite and tripartite relations can be seen as follows. To obtain the bipartite relation (167) from the tripartite relation (209), apply the latter to a purification $|\psi\rangle_{ABC}$ of ρ_{AB}. Now for tripartite pure states we have,

$$H(Z|C) = H(Z|B) - H(A|B)\,, \tag{211}$$

and inserting this into (209) gives (167).

Conversely we first prove (209) for tripartite pure states $|\psi\rangle_{ABC}$ by inserting (211) into (167). Then, note that the proof for mixed states ρ_{ABC} follows by applying (209) to a purification $|\psi\rangle_{ABCD}$ of ρ_{ABC}, and making use of the data-processing inequality (155),

$$H(Z|CD) \leq H(Z|C)\,. \tag{212}$$

The original proof of (209) was based on so-called smooth entropies.[18] The proof was subsequently simplified by Tomamichel and Renner (2011) and Coles *et al.* (2011), which culminated in the concise proof given by Coles *et al.* (2012). The latter proof distills the main ideas of the previous proofs: the use of duality relations for entropies as in (156) and the data-processing inequality as in (155). More generally, the proof technique applies to a whole family of entropies satisfying a few axioms (including the Rényi entropies). We will present the proof in App. C.3. Finally, we note that a direct matrix analysis proof was given by Frank and Lieb (2013a).

[18] We refer to (Tomamichel, 2016) for an introduction to smooth entropies.

3. Quantum memory tightens the bound

Here we argue that the tripartite uncertainty relation in terms of quantum memory (209) is tighter than the corresponding relation in terms of classical memory (166). We will show that there exists states ρ_{ABC} for which (209) is an equality but (166) is loose, even if one optimizes over all choices of measurements on B and C.

First, let us introduce some notation. Consider a bipartite state ρ_{AB} and let $\mathbb{X}_A$ and $\mathbb{Y}_B$ be measurements on systems A and B, respectively. Now, how small can we make the uncertainty $\mathbb{X}_A$ given that we can optimize over all choices of $\mathbb{Y}_B$? That is, consider the quantity

$$\alpha(\mathbb{X}_A, \rho_{AB}) := \min_{\mathbb{Y}_B} H(X_A|Y_B)\,. \tag{213}$$

This is to be compared to the classical-quantum conditional entropy

$$\beta(\mathbb{X}_A, \rho_{AB}) := H(X_A|B)\,. \tag{214}$$

Due to the data-processing inequality (155), we have that

$$\alpha(\mathbb{X}_A, \rho_{AB}) \geq \beta(\mathbb{X}_A, \rho_{AB})\,, \tag{215}$$

and naively one might guess that (215) is satisfied with equality in general. However, this is false (DiVincenzo *et al.*, 2004; Hiai and Petz, 1991). In general there is a non-zero gap:

$$\alpha - \beta > 0\,. \tag{216}$$

There are many examples to illustrate this, in fact one can argue that *most* states ρ_{AB} exhibit a gap between α and β (Dupuis *et al.*, 2013). This phenomenon is called locking and we discuss it more in Sec. VI.H.3. It is closely related to a measure of quantum correlations known as quantum discord (Modi *et al.*, 2012; Ollivier and Zurek, 2001): non-zero discord is associated with the potential to have a gap between α and β. We discuss discord in more detail in Sec. VI.H.2. For now we note that discord is defined as,

$$D(A|B) := \min_{\mathbb{Y}_B} H(A|Y_B) - H(A|B) \tag{217}$$

where the optimization is over all POVMs $\mathbb{Y}_B$ on B. Consider the following example.

Example 25. *Let* $\mathbb{X}_A = \{|0\rangle\langle 0|, |1\rangle\langle 1|\}$ *and consider the bipartite quantum state:*

$$\rho_{AB} = \frac{1}{2}\Big(|0\rangle\langle 0| \otimes |0\rangle\langle 0| + |1\rangle\langle 1| \otimes |+\rangle\langle +|\Big)\,. \tag{218}$$

For this state, the gap between α *and* β *is precisely given by the discord,*

$$D(A|B) = \alpha(\mathbb{X}_A, \rho_{AB}) - \beta(\mathbb{X}_A, \rho_{AB})\,. \tag{219}$$

It is known that $D(A|B) = 0$ *if and only if system* B *is classical, i.e., if* ρ_{AB} *is a quantum-classical state. But the state* ρ_{AB} *in* (218) *is not quantum-classical. Hence,* $D(A|B) > 0$ *and we have* $\alpha > \beta$.

Now we give an example state for which the quantum memory relation (209) is an equality but the measured relation (166) is loose.

Example 26. *Let* $d_A = d_B = d_C = 2$ *and consider the tripartite state* $|\psi\rangle_{ABC} = (|000\rangle + |11+\rangle)/\sqrt{2}$, *with* $\mathbb{Z}$ *being the standard basis and* $\mathbb{X}$ *being the* $\{|+\rangle, |-\rangle\}$ *basis. We have*

$$H(Z|C) = 1 - H(\rho_C) \approx 0.4 \tag{220}$$

$$H(X|B) = H(\rho_C) \approx 0.6\,. \tag{221}$$

Hence, this state satisfies the quantum memory relation (208) *with equality,*

$$H(X|B) + H(Z|C) = 1\,. \tag{222}$$

However, the classical memory relation (166) *is not satisfied with equality. This follows from Ex. 25, noting that* ρ_{AC} *is the same state as in* (218).

4. Tripartite guessing game

We note that also tripartite uncertainty relations can be understood in the language of guessing games as outlined in Figure 10, in particular, the concept of a monogamy game. Tomamichel *et al.* (2013) show that there is fundamental trade-off between Bob's guessing probability $p_{\text{guess}}(K|B\Theta)$ and Charlie's guessing probability $p_{\text{guess}}(K|C\Theta)$,

$$p_{\text{guess}}(K|B\Theta) + p_{\text{guess}}(K|C\Theta) \leq 2b\,, \tag{223}$$

with the overlap b as in (40). Alternatively, one can rewrite this in terms of the min-entropy using the concavity of the logarithm,

$$H_{\min}(K|B\Theta) + H_{\min}(K|C\Theta) \geq 2\log\frac{1}{b}\,. \tag{224}$$

Note that this relation (224) is an extension of (71) to the tripartite scenario. These relations again show a trade-off between Bob's and Charlie's winning probability, which is closely connected to the idea of monogamy of entanglement (cf. Sec. IV.E.1).

5. Extension to Rényi entropies

The Maassen-Uffink relation for Rényi entropies (35) naturally generalizes to a tripartite uncertainty relation with quantum memory. It is expressed in terms of the conditional Rényi entropies, whose definition and properties are discussed in App. C. For these entropies, the following relation holds (Coles *et al.*, 2012):[19]

$$H_\alpha(X|B) + H_\beta(Z|C) \geq \log\frac{1}{c} \quad \text{for} \quad \frac{1}{\alpha} + \frac{1}{\beta} = 2\,. \tag{225}$$

Notably, the tripartite uncertainty relation (209) is the special case where $\alpha = \beta = 1$. Another interesting special case is $\alpha = \infty$ and $\beta = 1/2$, which respectively correspond to the min- and max-entropies that we introduced in (140) and (157). The resulting relation,

$$H_{\min}(X|B) + H_{\max}(Z|C) \geq \log\frac{1}{c}\,, \tag{226}$$

was first proved by Tomamichel and Renner (2011), and is fundamental to quantum key distribution (see Sec. VI.B).

6. Arbitrary measurements

All of the tripartite uncertainty relations stated above can be generalized to arbitrary POVMs $\mathbb{X}$ and $\mathbb{Z}$. Coles *et al.* (2011) and Tomamichel and Renner (2011) independently noted that (209) holds for POVMs with the overlap c given by (49). This was strengthened by Tomamichel (2012) to the overlap c' given by (51). Further strengthening was given by Coles and Piani (2014b). However, their bound is implicit, involving an optimization over a single real-valued parameter over a bounded interval. Namely, they show a lower bound

$$q_{\text{CP2}} := \max_{0\leq p\leq 1} \lambda_{\min}(\Delta(p))\,, \tag{227}$$

where $\lambda_{\min}[\cdot]$ denotes the minimum eigenvalue and

$$\Delta(p) := p\,\delta(\mathbb{X},\mathbb{Z}) + (1-p)\delta(\mathbb{Z},\mathbb{X}) \tag{228}$$

$$\delta(\mathbb{X},\mathbb{Z}) := \sum_x a_x(\mathbb{X},\mathbb{Z}) \cdot \mathbb{X}^x \tag{229}$$

$$a_x(\mathbb{X},\mathbb{Z}) := -\log\Big\|\sum_z \mathbb{Z}^z\mathbb{X}^x\mathbb{Z}^z\Big\|\,. \tag{230}$$

Using the fact that $\delta(\mathbb{X},\mathbb{Z}) \geq \min_x a_x(\mathbb{X},\mathbb{Z}) \cdot \mathbb{1}$, it is straightforward to show that $q_{\text{CP2}} \geq \log(1/c')$.

[19] More precisely, the relation follows from the work (Coles *et al.*, 2012) in conjunction with properties of the conditional Rényi entropy presented in (Müller-Lennert *et al.*, 2013). It is thus first mentioned in the later work (Müller-Lennert *et al.*, 2013). Notably Coles *et al.* (2012) proves a tripartite uncertainty relation for a different definition of the conditional Rényi entropy (Tomamichel *et al.*, 2009).

F. Mutual information approach

While entropy quantifies the lack of information, it is both intuitive and useful to also consider measures that quantify the *presence* of information or correlation. Consider the *mutual information* $I(X:Y)$, which quantities the correlation between random variables X and Y, and is given by

$$I(X:Y) := H(X) + H(Y) - H(XY) \tag{231}$$
$$= H(X) - H(X|Y)\,. \tag{232}$$

Mutual information is a central quantity in information theory (Cover and Thomas, 1991; Shannon, 1948). It quantifies the information gained—or equivalently, the reduction of ignorance—about X when given access to Y.

1. Information exclusion principle

Hall pioneered an alternative formulation of the uncertainty principle based on the mutual information, which he called the *information exclusion principle* (Hall, 1995, 1997). Information exclusion relations are closely related to entropic uncertainty relations that allow for memory, hence is it natural to discuss them here. The idea is that one is interested in the trade-off between a memory system Y being correlated to $\mathbb{X}$ versus being correlated to $\mathbb{Z}$ (with $\mathbb{X}$ and $\mathbb{Z}$ being two measurements on some quantum system A).

It is worth noting that the mutual information is particularly well-suited for applications involving the transmission of information over a quantum channel. Here, one is interested in the correlations between a receiver and a sender. Hence, we will also discuss the application of information exclusion relations to transmission over channels.

2. Classical memory

Following Hall's derivation (Hall, 1995), we show now how information exclusion relations follow directly from entropic uncertainty relations. Consider a generic uncertainty relation involving Shannon entropy terms, of the form $\sum_{n=1}^{N} H(X_n)_\rho \geq q$ as in (159). Recall the discussion in Sec. IV.C which showed that the uncertainty relation $\sum_{n=1}^{N} H(X_n|Y) \geq q$ as in (162) immediately follows, where Y is some classical memory. Now with the definition of the mutual information (231) we can rewrite this as

$$\sum_{n=1}^{N} H(X_n) - I(X_n:Y) \geq q\,. \tag{233}$$

We have $H(X_n) \leq \log d$ for each n with d the dimension of the quantum system A that is measured. Combining this with (233) gives

$$\sum_{n=1}^{N} I(X_n:Y) \leq N \log d - q\,. \tag{234}$$

For example, if we take the Maassen-Uffink relation (31) as the starting point, we end up with

$$I(X:Y) + I(Z:Y) \leq \log(d^2 c) =: r_{\mathrm{H}} \tag{235}$$

The information exclusion relation in (235) was presented in (Hall, 1995). Note that we have $\log d \leq r_{\mathrm{H}} \leq 2\log d$, with r_{H} reaching the extreme points respectively for $c = 1/d$ and $c = 1$.

Equation (235) has an intuitive interpretation: any classical memory cannot be highly correlated to two complementary measurement outcomes of a quantum system. In the fully complementary case, the bound becomes $r_{\mathrm{H}} = \log d$, implying that if the classical memory is perfectly correlated to X, $I(X:Y) = \log d$, then it must be completely uncorrelated to Z, $I(Z:Y) = 0$.

3. Stronger bounds

Notice that (235) uses the same overlap c as appearing the Maassen-Uffink uncertainty relation (31). However, Grudka *et al.* (2013) realized that this often leads to a fairly weak bound. Rather the complementarity of the mutual information should depend highly on more elements of the overlap matrix $[c_{xz}]$ as in (32) than simply the maximum element c. They conjectured a stronger information exclusion relation, of the form

$$I(X:Y) + I(Z:Y) \leq r_{\mathrm{G}}, \quad r_{\mathrm{G}} = \log_2\left(d \cdot \sum_{\text{d largest}} c_{xz}\right), \tag{236}$$

with the sum over the largest d terms of the matrix $[c_{xz}]$. This conjecture was proved in (Coles and Piani, 2014b), where the bound was further strengthened to

$$I(X:Y) + I(Z:Y) \leq r_{\mathrm{CP}}, \tag{237}$$

with

$$r_{\mathrm{CP}} = \min\left\{r(\mathbb{X},\mathbb{Z}), r(\mathbb{Z},\mathbb{X})\right\} \tag{238a}$$
$$r(\mathbb{X},\mathbb{Z}) = \log\left(d \sum_x \max_z c_{xz}\right) \tag{238b}$$
$$r(\mathbb{Z},\mathbb{X}) = \log\left(d \sum_z \max_x c_{xz}\right)\,. \tag{238c}$$

One can easily verify that $r_{\mathrm{CP}} \leq r_{\mathrm{G}} \leq r_{\mathrm{H}}$.

Example 27. *The unitary in* (43) *from Ex.* 8 *provides a simple example where all three bounds are different, namely* $r_{\rm H} = \log 6$, $r_{\rm G} = \log 5$, *and* $r_{\rm CP} = \log(9/2)$.

Notice that the behavior of the bounds $r_{\rm H}$ and $r_{\rm CP}$ are qualitatively different in that they become trivial under different conditions. The former is trivial if at least one row or column of $[c_{xz}]$ is trivial (i.e., composed of all zeros except for one element being one), whereas the latter is trivial only if all rows and columns $[c_{xz}]$ are trivial. Hence, the latter gives a non-trivial bound for a much larger range of scenarios.

4. Quantum memory

It is natural to ask whether system Y can be generalized to a quantum memory B. Coles and Piani (2014b) show that (237) can indeed be generalized to

$$I(X:B) + I(Z:B) \leq r_{\rm CP} - H(A|B)\,. \tag{239}$$

Here the quantum mutual information of a bipartite quantum state ρ_{AB} is defined as

$$I(A:B) := H(\rho_A) + H(\rho_B) - H(\rho_{AB}) \tag{240}$$
$$= H(\rho_A) - H(A|B)\,, \tag{241}$$

and evaluated on the classical-quantum state ρ_{XB} as in (168). Notice that if we specialize to the case where $B = Y$ is classical, then $H(A|Y) \geq 0$ and hence (239) also implies (237).

Example 28. *Consider the extreme case where* ρ_{AB} *is maximally entangled. In this case, both* $I(X:B)$ *and* $I(Z:B)$ *are maximal, each being equal to* $\log d$. *Hence, any upper bound on them must be trivial and bounding their sum by* $r_{\rm CP}$ *would not work. The bound must be weakened in such a way that it becomes trivial, and indeed the term* $-H(A|B)$ *accomplishes this. Namely, we have* $-H(A|B) = \log d$ *in the maximally entangled case.*

In general, a negative value of $H(A|B)$ implies that ρ_{AB} has distillable entanglement, and this results in a bound in (239) that is larger than $r_{\rm CP}$. In the other extreme, when $H(A|B)$ is positive, which intuitively means that the correlations between Alice and Bob are weak, (239) strengthens the bound in (237).

5. A conjecture

Following the resolved conjectures by Grudka *et al.* (2013); Kraus (1987); and Renes and Boileau (2009), we point to some recent and open conjecture by Schneeloch *et al.* (2014). They ask if for any bipartite quantum state ρ_{AB},

$$I(X_A:X_B) + I(Z_A:Z_B) \overset{?}{\leq} I(A:B)\,, \tag{242}$$

where X_A and Z_A are the registers associated with measuring two MUBs $\mathbb{X}_A$ and $\mathbb{Z}_A$ on system A, and likewise for X_B and Z_B on system B. The relation (242) would say that the quantum mutual information is lower bounded by the sum of the classical mutual informations in two mutually unbiased bases. We note that a stronger conjecture, in which X_B and Z_B are replaced by the quantum memory B, can be violated in general.

G. Quantum channel formulation

1. Bipartite formulation

Christandl and Winter (2005) considered the question of how well information can be transmitted over a quantum channel. A quantum channel is the general form for quantum dynamics (more general than unitary evolution). Mathematically a quantum channel $\mathcal{E}$ is a completely positive trace preserving map, and can be represented in its Kraus form,

$$\mathcal{E}(\cdot) = \sum_j K_j(\cdot)K_j^\dagger\,, \quad \text{where} \sum_j K_j^\dagger K_j = \mathbb{1}\,. \tag{243}$$

Christandl and Winter (2005) address the topic of sending classical information over a quantum channel, or more specifically, sending two complementary types of classical information over a quantum channel. They consider a scenario where Alice chooses a state, with probability $1/d$, from a set of d orthonormal states, which we label as $\mathbb{Z} = \{|\mathbb{Z}^z\rangle\langle\mathbb{Z}^z|\}$. She then sends the state over the channel $\mathcal{E}$ to Bob, and Bob tries to distinguish which state she sent. Likewise Alice and Bob may play the same game but with the $\mathbb{X} = \{|\mathbb{X}^x\rangle\langle\mathbb{X}^x|\}$ states instead, where the $\mathbb{X}$ and $\mathbb{Z}$ states are related by the Fourier matrix F, given by (207). Bob's distinguishability for the $\mathbb{Z}$ states can be quantified by the so-called Holevo quantity (Holevo, 1973),

$$\chi(\mathcal{E},\mathbb{Z}) = H\left(\sum_z \frac{1}{d}\mathcal{E}\Big(|\mathbb{Z}^z\rangle\langle\mathbb{Z}^z|\Big)\right) - \sum_z \frac{1}{d} H\Big(\mathcal{E}\big(|\mathbb{Z}^z\rangle\langle\mathbb{Z}^z|\big)\Big)\,. \tag{244}$$

Likewise, $\chi(\mathcal{E},\mathbb{X})$ is a measure of Bob's distinguishability for the $\mathbb{X}$ states, and is defined analogously to (244). Christandl and Winter (2005) prove that

$$\chi(\mathcal{E},\mathbb{X}) + \chi(\mathcal{E},\mathbb{Z}) \leq \log d + I_{\rm coh}\left(\frac{\mathbb{1}}{d},\mathcal{E}\right)\,, \tag{245}$$

where the coherent information $I_{\rm coh}(\rho,\mathcal{E})$ is a measure of the quality of a quantum channel $\mathcal{E}$ introduced by Schumacher and Nielsen (1996). For the maximally-mixed

input state $\mathbb{1}/d$ it is given by

$$I_{\text{coh}}\left(\frac{\mathbb{1}}{d},\mathcal{E}\right) = H\Big(\mathcal{E}(\mathbb{1}/d)\Big) - H\Big((\mathcal{I}\otimes\mathcal{E})\big(|\Phi\rangle\langle\Phi|\big)\Big)\,, \tag{246}$$

where $|\Phi\rangle = \sum_j (1/\sqrt{d})|j\rangle|j\rangle$ is a maximally entangled state. Coles *et al.* (2011) noted that (245) holds for arbitrary MUBs, and that it naturally generalizes to arbitrary orthonormal bases $\mathbb{X}$ and $\mathbb{Z}$ with the right-hand side of (245) replaced by

$$\log\left(d^2 c\right) + I_{\text{coh}}\left(\frac{\mathbb{1}}{d},\mathcal{E}\right)\,. \tag{247}$$

Later this bound was improved by Coles and Piani (2014b) to

$$r_{\text{CP}} + I_{\text{coh}}\left(\frac{\mathbb{1}}{d},\mathcal{E}\right)\,. \tag{248}$$

While (245) may look similar to some uncertainty relations discussed previously in this section, especially (239), it is important to note the conceptual difference. The relations discussed previously were from a static perspective, whereas (244) refers to a dynamic perspective involving a sender and a receiver. Intuitively, what (245) says is that if Alice can transmit both the $\mathbb{Z}$ states and the $\mathbb{X}$ states well to Bob, then $\mathcal{E}$ is a noiseless quantum channel, i.e., it is close to a perfect channel, as quantified by the coherent information.

2. Static-dynamic isomorphism

With that said, there is a close, mathematical relationship between the static and dynamic perspectives. In fact, there is an isomorphism, known as the Choi-Jamiołkowski isomorphism (Choi, 1975; Jamiołkowski, 1972), that relates the two perspectives (e.g., see (Życzkowski and Bengtsson, 2004)). Every quantum channel $\mathcal{E}$ corresponds to a bipartite mixed state defined by,

$$\rho_{AB} = (\mathcal{I}\otimes\mathcal{E})(|\Phi\rangle\langle\Phi|)\,, \tag{249}$$

where $|\Phi\rangle = \sum_j (1/\sqrt{d})|j\rangle|j\rangle$ is maximally entangled (see Fig. 11(a)). Note that ρ_{AB} here has the property that $\rho_A = \text{tr}_B(\rho_{AB}) = \mathbb{1}/d_A$ is maximally mixed. Likewise, every bipartite mixed state ρ_{AB} with marginal $\rho_A = \mathbb{1}/d_A$ corresponds to a quantum channel whose action on some operator O is defined as,

$$\mathcal{E}(O) = d_A\,\text{tr}_A\left[\left(O^T\otimes\mathbb{1}\right)\rho_{AB}\right], \tag{250}$$

where the transpose denoted by $(\cdot)^T$ is taken in the standard basis. One can easily verify that the condition that $\rho_A = \mathbb{1}/d_A$ is connected to the fact that $\mathcal{E}$ is trace-preserving.

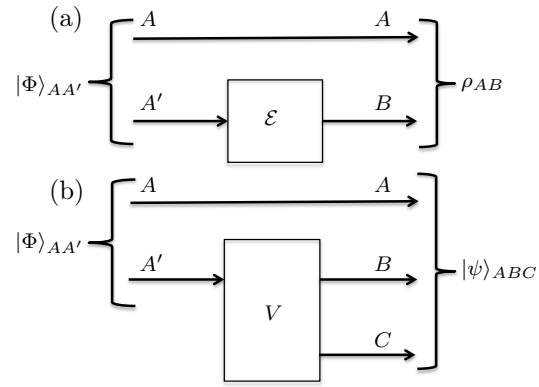


FIG. 11 How to convert the dynamical evolution of a system into (a) a bipartite mixed state or (b) a tripartite pure state.

This isomorphism can be exploited to derive uncertainty relations for quantum channels as corollaries from uncertainty relations for states, and vice versa. This point was emphasized, e.g., in (Coles *et al.*, 2011). For example, if one has an uncertainty relation for bipartite states ρ_{AB}, such as (167), then one can apply this relation to the state in (249) in order to obtain an uncertainty relation for channels.

Specifically, notice that if Alice measures observable $\mathbb{Z}$ on system A in Fig. 11(a) and obtains outcome $|\mathbb{Z}^z\rangle\langle\mathbb{Z}^z|$, then the state corresponding to the transpose, $|\mathbb{Z}^z\rangle\langle\mathbb{Z}^z|^T$, will be sent through the channel $\mathcal{E}$. This implies that the Holevo quantity can be thought of as a classical-quantum mutual information as,

$$\chi(\mathcal{E},\mathbb{Z}^T) = I(Z:B)_\rho = \log d - H(Z|B)_\rho \quad \text{where} \quad \mathbb{Z}^T = \left\{|\mathbb{Z}^z\rangle\langle\mathbb{Z}^z|^T\right\}, \tag{251}$$

and the right-hand side is evaluated for the state

$$\rho_{ZB} = \sum_z |z\rangle\langle z|\otimes\text{tr}_A\left[(|\mathbb{Z}^z\rangle\langle\mathbb{Z}^z|\otimes\mathbb{1}_B)\rho_{AB}\right] \tag{252}$$

$$= \sum_z \frac{1}{d}|z\rangle\langle z|\otimes\mathcal{E}\left(|\mathbb{Z}^z\rangle\langle\mathbb{Z}^z|^T\right)\,. \tag{253}$$

Using (251), one can verify that the channel uncertainty relation (245) is a corollary of the bipartite state uncertainty relation, either (167) or (239).

3. Tripartite formulation

One can formulate uncertainty relations for a dynamic tripartite scenario where Alice sends the $\mathbb{Z}$ states over quantum channel $\mathcal{E}$ to Bob or the $\mathbb{X}$ states over the *complementary* quantum channel $\mathcal{F}$ to Charlie. The relationship between a channel and its complementary channel can be seen via the Stinespring dilation (Stinespring, 1955), in which one writes the channel in terms of an

isometry V that maps $\mathcal{H}_A \to \mathcal{H}_{BC}$, namely

$$\mathcal{E}(O) = \mathrm{tr}_C[VOV^\dagger], \tag{254}$$

$$\mathcal{F}(O) = \mathrm{tr}_B[VOV^\dagger]\,. \tag{255}$$

Analogous to (249), we consider the tripartite pure state defined by

$$|\psi\rangle_{ABC} = (\mathbb{1} \otimes V)|\Phi\rangle\,. \tag{256}$$

This mapping is depicted in Fig. 11(b). The tripartite uncertainty relations presented in Section IV.E can then be applied to the state $|\psi\rangle_{ABC}$ in (256) in order to derive uncertainty relations for complementary quantum channels. For example, Coles *et al.* (2011) read (209) in this sense to obtain

$$\chi(\mathcal{E}, \mathbb{X}) + \chi(\mathcal{F}, \mathbb{Z}) \leq \log\left(d^2 c\right)\,, \tag{257}$$

for two orthonormal bases $\mathbb{X}$ and $\mathbb{Z}$. This relation implies that if Alice can send the $\mathbb{Z}$ states well to Charlie over the channel $\mathcal{F}$, then Bob cannot distinguish very well the outputs of the channel $\mathcal{E}$ associated with Alice sending a complementary set of states $\mathbb{X}$.

V. POSITION-MOMENTUM UNCERTAINTY RELATIONS

As discussed in the introduction (Sec. I), the first precise statement of the uncertainty principle was formulated for position and momentum measurements. Namely, Kennard (1927) has shown that for all states (setting $\hbar = 1$),

$$\sigma(Q) \cdot \sigma(P) \geq \frac{1}{2}\,, \tag{258}$$

where $\sigma(Q)$ denotes the standard deviation of the probability density $\Gamma_Q(q)$ when measuring the position Q, and similarly for $\sigma(P)$ when measuring the momentum P. Henceforth, we always set $\hbar = 1$.

Example 29. *Consider Gaussian wave packets with position probability density (see also Fig. 12)*

$$\Gamma_Q(q) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot \exp\left(-q^2 \cdot \frac{1}{2\sigma^2}\right)\,, \tag{259}$$

and corresponding momentum probability density

$$\Gamma_P(p) = \sqrt{\frac{2\sigma^2}{\pi}} \cdot \exp\left(-p^2 \cdot 2\sigma^2\right)\,. \tag{260}$$

It is then straightforward to check that these achieve equality in (258) and hence minimize the uncertainty in terms of the product of the two standard deviations.

In contrast to Kennard's formulation (258) of the uncertainty principle the relations developed in Sec. III–IV are phrased in terms of entropy measures and only apply to finite-dimensional systems (whereas position and momentum measurements can only be modeled on infinite-dimensional spaces). In this section we review entropic uncertainty relations with and without a memory system for position and momentum measurements.[20] We discuss applications to continuous variable quantum cryptography later in Sec. VI.B.5.

A. Entropy for infinite-dimensional systems

On a technical level, the position operator Q and the momentum operator P with the canonical commutation relation

$$[P, Q] = i\mathbb{1} \tag{261}$$

can only be represented as unbounded operators on infinite-dimensional spaces. Hence, we need to extend our setup from finite dimensional Hilbert spaces to separable Hilbert spaces A with $\dim(A) = \infty$. However, quantum states can still be represented as linear, positive semi-definite operators. Hence, we just keep the notation the same as for finite-dimensional spaces without going into any mathematical details. We start with describing how to define entropy for infinite-dimensional systems.

1. Shannon entropy for discrete distributions

Imagine a finite resolution detector that measures the position Q by indicating in which interval

$$\mathcal{I}_{k;\delta} := \big(k\delta, (k+1)\delta\big] \qquad (k \in \mathbb{Z})\,, \tag{262}$$

of size $\delta > 0$ the value q falls. This defines a discrete probability distribution Γ_{Q_δ} with infinitely many elements.[21] If the initial is described by a pure state wave function $|\psi(q)\rangle_Q$ we get $\{\Gamma_{Q_\delta}(k)\}_{k\in\mathbb{Z}}$ with

$$\Gamma_{Q_\delta}(k) = \int_{k\delta}^{(k+1)\delta} \big|\langle\psi(q)|\psi(q)\rangle_Q\big|\,\mathrm{d}q\,. \tag{263}$$

We then just define the Shannon entropy of Γ_{Q_δ} in the usual way as

$$H(Q_\delta) := -\sum_{k=-\infty}^{\infty} \Gamma_{Q_\delta}(k) \log \Gamma_{Q_\delta}(k)\,. \tag{264}$$

---

[20] Entropic uncertainty relations for completely general quantum systems described by von Neumann algebras and measurements described by measure spaces are also studied in the literature (Frank and Lieb, 2013a; Furrer *et al.*, 2014). However, here we focus on the physically relevant position and momentum measurements.

[21] In all of Sec. V, we use the letter Γ instead of P for probability distributions since the momentum operator is already denoted by the letter P.

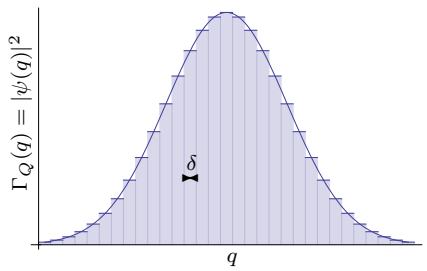


FIG. 12 Gaussian wave packet in position space with $\Gamma_Q(q)$ as in (259). We also show the finite resolution discretization from (263) in intervals of size δ.

Despite the fact that there are now infinitely many terms in the sum, $H(Q_\delta)$ keeps many of the properties of its finite-dimensional analogue. In particular, $H(Q_\delta) \geq 0$ and the Shannon entropy can still be thought of as an information measure.

2. Shannon entropy for continuous distributions

The differential Shannon entropy is defined in the limit of infinitely small interval size $\delta \to 0$,

$$h(Q) := \lim_{\delta\to 0}\Big(H(Q_\delta) + \log\delta\Big) \tag{265}$$

$$= \lim_{\delta\to 0}\left(-\sum_{k=-\infty}^{\infty}\Gamma_{Q_\delta}(k)\log\frac{\Gamma_{Q_\delta}(k)}{\delta}\right). \tag{266}$$

The term $H(Q_\delta)$ scales with the interval $\delta \to 0$ and hence the normalization in (265). This makes the differential Shannon entropy an entropy density in the limit $\delta \to 0$. We also get a closed formula for the differential Shannon entropy (at least when $\Gamma_Q(q)$ is continuous),

$$h(Q) = -\int \mathrm{d}q\,\Gamma_Q(q)\log\Gamma_Q(q)\,, \tag{267}$$

where $\Gamma_Q(q)$ denotes the probability density when measuring the position Q. For the momentum probability density $\Gamma_P(p)$ we define the discrete and differential Shannon entropy in the same way. Since probability densities can be larger than one, not all of the properties of discrete Shannon entropy carry over. For example the differential Shannon entropy can be negative.

Example 30. *For Gaussian wave packets as in* (259) *and* (260) *we have,*

$$h(Q) = \frac{1}{2}\log\left(2\pi e\sigma^2\right) \quad \text{and} \quad h(P) = \frac{1}{2}\log\frac{\pi e}{2\sigma^2}\,. \tag{268}$$

By inspection we find that $h(Q) < 0$ for σ sufficiently small and $h(P) < 0$ for σ sufficiently large.

Nevertheless the uncertainty principle can still be expressed in term of differential Shannon entropies.

B. Differential relations

Extending the work of Everett (1957) and Hirschman (1957), Białynicki-Birula and Mycielski (1975) and independently Beckner (1975) show for position and momentum measurements Q and P, respectively, that

$$h(P) + h(Q) \geq \log(e\pi)\,. \tag{269}$$

We emphasize that (269) holds even though each of the differential Shannon entropies on the left-hand side can become negative! As in Kennard's relation (258) Gaussian wave packets again minimize the uncertainty and lead to equality in (269). This shows that the relation is tight. As sketched in Sec. II the entropic relation (269) also implies Kennard's relation (258) and is therefore strictly more general.

Recently an alternative bound was shown by Frank and Lieb (2012), see also the works of Hall and Wiseman (2012) and Rumin (2012). Extending the work of Beckner (1975), Hall (1999), and Rumin (2011), Frank and Lieb show that

$$h(Q) + h(P) \geq \log(2\pi) + H(\rho_A)\,, \tag{270}$$

where

$$H(\rho_A) := -\operatorname{tr}\left[\rho_A\log\rho_A\right] \tag{271}$$

denotes the von Neumann entropy of the infinite-dimensional input state before any measurement was performed. We note that in contrast to the differential Shannon entropies on the left-hand side, the von Neumann entropy is always non-negative (even for infinite-dimensional systems). The state independent bound $\log(2\pi) \leq \log(e\pi)$ is worse, but interestingly (270) becomes an equality for a thermal state in the infinite temperature limit (Frank and Lieb, 2012; Hall, 1999). Hence, the relation (270) is also tight if we insist on having the von Neumann entropy $H(A)_\rho$ on the right-hand side.

C. Finite-spacing relations

It has been argued in the literature that ideal position and momentum measurements can effectively never be performed because every detector has a finite accuracy. We can then ask: in what other than a purely mathematical sense does (269) and (270) express the uncertainty principle?[22] Certainly a more operational way to express uncertainty is in terms of the discrete Shannon entropy as defined in (264). A series of works (Białynicki-Birula, 1984; Partovi, 1983; Rudnicki, 2011; Rudnicki

[22] This criticism also applies to Kennard's relation (258) and a finite spacing version thereof has been derived in (Rudnicki *et al.*, 2012).

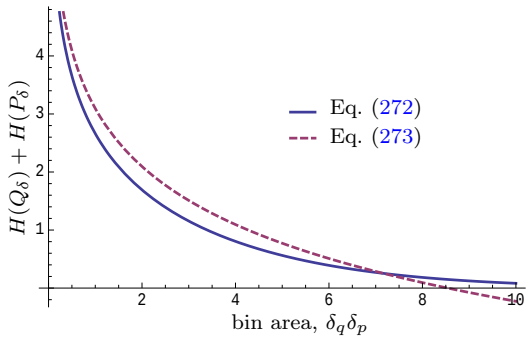


FIG. 13 Comparison of the lower bounds in (272) and (273) on the uncertainty generated by the finite-spacing position and momentum measurements Q_δ, P_δ as in (263). Note that the latter bound becomes negative and hence trivial for larger spacings $\delta_q\delta_p \gtrsim 8.5$.

et al., 2012) established that for measurements with finite spacing δ_q for the position q and finite spacing δ_p for the momentum we have that

$$H(Q_\delta) + H(P_\delta) \geq \log(2\pi) - \log\left(\delta_q\delta_p \cdot S_0^{(1)}\left(1, \frac{\delta_q\delta_p}{4}\right)^2\right), \tag{272}$$

where $S_0^{(1)}(\cdot,\cdot)$ denotes the 0th radial prolate spheroidal wave function of the first kind (Slepian and Pollak, 1961). This way of expressing the uncertainty principle has the advantage that the discrete Shannon entropy is always non-negative and has a clear information theoretic interpretation. As we will say later, it is the discrete formulation of the uncertainty principle that becomes relevant for applications in continuous variable quantum cryptography (see Sec. V.E and VI.B.5).

Interestingly (272) is not tight for general $\delta > 0$ since we also know that (Białynicki-Birula, 1984),

$$H(Q_{\delta q}) + H(P_{\delta p}) \geq \log(e\pi) - \log\left(\delta_q\delta_p\right), \tag{273}$$

which becomes tighter for $\delta \to 0$ (see Fig. 13). Rudnicki (2015) employs a majorization based approach along the lines of Sec. III.I to improve on (272) and (273) for large spacing. However, this does not yield a nice closed formula and we refer to Rudnicki (2011, 2015) for a discussion of tightness and a more detailed comparison. We will further comment on this issue in the next section (Sec. V.D) after extending (269) and (272) to a quantum memory system.

D. Uncertainty given a memory system

For finite-dimensional systems we can write the conditional von Neumman entropy of bipartite quantum states ρ_{AB} as $H(A|B) = H(AB) - H(B)$. However, for infinite-dimensional systems this is not a sensible notion of conditional entropy. This is because for some states both terms $H(AB)$ and $H(B)$ can become infinite even though the conditional entropy is finite.

Example 31. *Consider a bipartite system with A a qubit and B infinite-dimensional. Let ρ_{AB} be a state that is entangled between A and a two-dimensional subspace of B, and with the marginal ρ_B having eigenvalues (Wehrl, 1978),*

$$\lambda_k \propto \frac{1}{k(\log k)^2} \quad \text{for} \quad k > 2\,. \tag{274}$$

Then, we have $H(AB) = \infty$ as well as $H(B) = \infty$ even though any sensible definition of conditional entropy $H(A|B)$ should be upper and lower bounded by plus and minus one, respectively.

However, observe that the classical quantum entropy of finite-dimensional classical quantum states ρ_{XB} as in (138) can be rewritten in terms of Umegaki's relative entropy (Umegaki, 1962),

$$D(\rho\|\sigma) := \mathrm{tr}[\rho(\log\rho - \log\sigma)] \tag{275}$$

$$\text{as} \quad H(X|B) = -\sum_x D(P_X(x)\rho_B^x\|\rho_B)\,. \tag{276}$$

Furrer *et al.* (2014) point out that (276) can be lifted to

$$H(Q_\delta|B) := -\sum_{k=-\infty}^{\infty} D\big(\rho_B^{k;\delta}\big\|\rho_B\big)\,, \tag{277}$$

where $\rho_B^{k;\delta}$ denotes the (sub-normalized) marginal state on B when the position Q is measured in $\mathcal{I}_{k;\delta}$, i.e.,

$$P_{Q_\delta}(k) := \mathrm{tr}\left[\rho_B^{k;\delta}\right] \text{ is the probability to measure in } \mathcal{I}_{k;\delta}. \tag{278}$$

1. Tripartite quantum memory uncertainty relations

With (277) as the definition for classical quantum entropy Furrer *et al.* (2014) find,

$$H(Q_{\delta q}|B) + H(P_{\delta p}|C) \geq \log(2\pi) - \log\left(\delta_q\delta_p \cdot S_0^{(1)}\left(1, \frac{\delta_q\delta_p}{4}\right)^2\right). \tag{279}$$

This is the extension of (272) to quantum memories and likewise not tight. By taking the limit $\delta \to 0$ we find the differential quantum conditional entropy

$$h(Q|B) := \lim_{\delta\to 0}\Big(H(Q_\delta|B) + \log\delta\Big) \tag{280}$$

$$= \int \mathrm{d}q\; D(\rho_B^q\|\rho_B)\,, \tag{281}$$

where the second equality holds under some finiteness assumption (Furrer *et al.*, 2014). With (279) we then immediately find the extension of (269) to quantum memories,

$$h(Q|B) + h(P|C) \geq \log(2\pi)\,. \tag{282}$$

Example 32. *For the EPR state on AB (or likewise AC) in the limit of perfect correlations (282) becomes an equality. For finite squeezing strength* $r = \operatorname{arccosh}(\nu)/2$ *the EPR state is a Gaussian state with covariance matrix*

$$\Gamma_{AB}(\nu) = \frac{1}{2}\begin{pmatrix} \nu\mathbb{1}_2 & \sqrt{\nu^2-1}Z_2 \\ \sqrt{\nu^2-1}Z_2 & \nu\mathbb{1}_2 \end{pmatrix}$$
$$\text{with } \mathbb{1}_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \text{ and } Z_2 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{283}$$

The left-hand side of (282) for the EPR state generated by $\Gamma_{AB}(\nu)$ *is then calculated to (Furrer* et al.*, 2014),*

$$h(Q|B) + h(P) = \log(e\pi\nu) - \frac{\nu+1}{2}\log\left(\frac{\nu+1}{2}\right) + \frac{\nu-1}{2}\log\left(\frac{\nu-1}{2}\right), \tag{284}$$

which converges to $\log(2\pi)$ *for* $\nu \to \infty$*. In Fig. 14 we plot (284) in terms of the squeezing strength* $r = \operatorname{arccosh}(\nu)/2$*:*

1. *For* $r = 0$ *the system B is uncorrelated and we have the lower bound* $h(Q)+h(P) \geq \log(e\pi)$ *as in (269).*

2. *For* $r > 0$ *we have to take the quantum memory B into account and only the lower bound* $h(Q|B) + h(P) \geq \log(2\pi)$ *from (282) holds.*

3. *For* $r \to \infty$ *we get maximal correlations and the bound (282) becomes an equality.*

We note that in typical experiments for applications (see Sec. VI.B.5) a squeezing strength of $r \approx 1.5$ *is achievable (Eberle* et al.*, 2013). For this the lower bound (282) is already very tight.*

Note that in (282) the state independent bound is $\log(2\pi)$ whereas it was $\log(e\pi)$ for the case without quantum memory in (269). Hence, in contrast to the finite-dimensional case, a quantum memory reduces the state independent uncertainty limit. This is because for the approximate EPR state there exists a gap between the accessible classical correlation and the classical-quantum correlation. That is, even when minimized over all measurements Q_B on B,

$$h(Q|Q_B) - h(Q|B) \approx \log(e/2)\,. \tag{285}$$

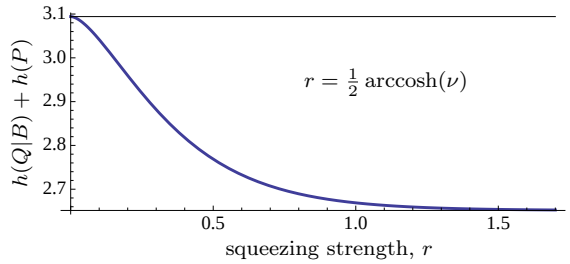


FIG. 14 We picture the uncertainty, $h(Q|B) + h(P)$, of the EPR state from Ex. 32 in terms of the squeezing strength r.

2. Bipartite quantum memory uncertainty relations

Similarly as for finite-dimensional systems (Sec. IV.D) it is possible to reformulate uncertainty relations with quantum memory in a bipartite form. For continuous position and momentum measurements Frank and Lieb (2013a) show that,

$$h(Q|B) + h(P|B) \geq \log(2\pi) + H(A|B)_\rho\,. \tag{286}$$

This is the extension of (270) to a quantum memory system. However, we note that (286) only holds if that all the terms appearing in $H(A|B) = H(AB) - H(B)$ are finite (which is in general too restrictive).[23]

3. Mutual information approach

Finally, a conceptually different approach was taken by Hall (1995) where the uncertainty relative to a memory system is quantified in terms mutual information instead of conditional entropy (see Sec. IV.F for a general discussion). Similarly as for the conditional entropy in (277), mutual information for classical quantum states is most generally defined in terms of Umegaki's relative entropy from (275),

$$I(Q_\delta : B) := \sum_{k=-\infty}^{\infty}\left(D\left(\rho_B^{k;\delta}\middle\|\rho_B\right) + H(B)_{\rho^{k;\delta}}\right). \tag{287}$$

In contrast to entropy, however, the mutual information stays finite when taking the limit $\delta \to 0$,

$$I(Q : B) := \lim_{\delta\to 0} I(Q_\delta : B)\,. \tag{288}$$

Hence, no regularization in terms of the interval size δ is taken. For classical memories M it was shown that (Hall, 1995),

$$I(Q : M) + I(P : M) \leq 1 + \log\sigma_\Gamma(Q) + \log\sigma_\Gamma(P)\,. \tag{289}$$

[23] This restriction is connected with the question about a sensible notion of conditional entropy for fully quantum states (Kuznetsova, 2011).

It is an open question to find a generalization that also holds for quantum memories. This would be in analogy to what is known for finite-dimensional systems (see Sec. IV.F.4).

E. Extension to min- and max-entropy

As for finite-dimensional systems (see Sec. III.C.2), entropic uncertainty relations like (269) and (272) can not only be shown for the Shannon entropy, but also more generally for pairs of Rényi entropies (Białynicki-Birula, 2006, 2007; Rudnicki *et al.*, 2012). Here we focus on a special case that is important for applications in continuous variable quantum cryptography (see Sec. VI.B.5): we study relations in terms of the Rényi entropy of order ∞ and its dual quantity the Rényi entropy of order $1/2$. These are exactly the min- and max-entropy, respectively.

1. Finite-spacing relations

Following the finite resolution detector picture as in (262) and (263), the conditional min-entropy is defined as

$$H_{\min}(Q_\delta|B) := -\log p_{\text{guess}}(Q_\delta|B)\,. \tag{290}$$

Here we have the optimal guessing probability as in (139),

$$p_{\text{guess}}(X|B) := \sup_{\mathbb{X}_B}\left\{\sum_{k=-\infty}^{\infty}\Gamma_{Q_\delta}(k)\,\mathrm{tr}\left[\mathbb{X}_B^k\rho_B^{k;\delta}\right] : \mathbb{X}_B \text{ POVM on } B\right\}. \tag{291}$$

In analogy to the finite-dimensional case, the min-entropy quantifies the uncertainty of the classical register Q_δ from the perspective of an observer with access to the quantum memory B. The conditional max-entropy is given by

$$H_{\max}(Q_\delta|B) := \log F_{\text{dec}}(Q_\delta|B)\,, \tag{292}$$

where we have the optimal decoupling fidelity

$$F_{\text{dec}}(Q_\delta|B) := \sup\left\{\left(\sum_{k=-\infty}^{\infty}\sqrt{F\left(\rho_B^{k;\delta},\sigma_B\right)}\right)^2 : \sigma_B \text{ state on } B\right\}. \tag{293}$$

The decoupling fidelity is a measure of how much information the quantum memory B contains about the classical register Q_δ.[24] For these definitions Furrer *et al.* (2014) show

$$H_{\min}(Q_\delta|B) + H_{\max}(P_\delta|C) \geq \log(2\pi) - \log\left(\delta_q\delta_p\cdot S_0^{(1)}\left(1,\frac{\delta_q\delta_p}{4}\right)^2\right)\,, \tag{294}$$

as well as the same relation with Q_δ and P_δ interchanged. We note that the special case with trivial quantum memories B, C was already shown in (Rudnicki *et al.*, 2012). Furrer *et al.* (2014) show that the relation (294) is tight for any spacing $\delta > 0$ even in the absence of any correlations (i.e., there exist states for which the relation becomes an equality). Note that this is contrast to the situation for the Shannon entropy, where neither (272) and (273), nor (279) are tight.

2. Differential relations

For the differential version we take the limit $\delta \to 0$,

$$h_{\min}(Q|B) := \lim_{\delta\to 0}\left(H_{\min}(Q_\delta|B) + \log\delta\right)\,, \tag{295}$$

and similarly for $h_{\max}(Q|B)$.[25] We then find the uncertainty relation (Furrer *et al.*, 2014),

$$h_{\min}(Q|B) + h_{\max}(P|C) \geq \log(2\pi) \tag{298}$$

as well as the same relation with Q and P interchanged. (Białynicki-Birula, 2006) shows that (298) becomes an equality for pure Gaussian states as in (259) and (260). Note that this implies in particular that the unconditional special case

$$h_{\min}(Q) + h_{\max}(P) \geq \log(2\pi) \tag{299}$$

is tight. Hence, the optimal state independent constant is $\log(2\pi)$ for the min- and max-entropy where as the optimal constant for the Shannon entropy in (269) is $\log(e\pi)$.

F. Other infinite-dimensional measurements

As a multidimensional extension of (269), Huang (2011) shows that for any measurements of the form

$$A = \sum_{i=1}^n a_iQ_i + a_i'P_i,\quad B = \sum_{i=1}^n b_iQ_i + b_i'P_i \quad \text{with} \quad a_i, a_i', b_i, b_i' \in \mathbb{R}\,, \tag{300}$$

---

[24] For finite-dimensional systems the expression (292) is equivalent to the max-entropy as defined in (157), see (Furrer *et al.*, 2014; König *et al.*, 2009).

[25] Under some finiteness assumptions we also have

$$h_{\max}(Q|B) = \log\sup\left\{\int \mathrm{d}q\,\left(\sqrt{F\left(\rho_B^q,\sigma_B\right)}\right)^2 : \sigma_B \text{ state on } B\right\} \tag{296}$$

$$h_{\min}(Q|B) = -\log\sup\left\{\int \mathrm{d}q\,\rho_B^q(\mathbb{X}_B^q) : q\mapsto\mathbb{X}_B^q \text{ POVM on } B\right\} \tag{297}$$

we have that

$$h(A) + h(B) \geq \log(e\pi) + \log|[A, B]| \,. \tag{301}$$

Huang (2011) also shows that for any measurement pair A, B as in (300) there exist states for which (301) becomes an equality.

Moreover, the techniques for deriving position-momentum uncertainty relations can also be applied to other complementary observable pairs that are modeled on infinite-dimensional Hilbert spaces. For example, for a particle on a circle we have the position angle φ and the conjugate angular momentum observable L_z. We consider a measurement device that either tells in which of

$$M := 2\pi/\delta_\varphi \quad \text{bins of size } \delta_\varphi \tag{302}$$

the particle is in or the exact value of the angular momentum L_z. Improving on the earlier earlier work of Partovi (1983), Białynicki-Birula (1984) shows that

$$H(\varphi_\delta) + H(L_z) \geq \log M \,, \tag{303}$$

where the discrete angle distribution $\tilde{P}_{\varphi_\delta}$ is defined similarly as in (263), and $\tilde{P}_{L_z}$ denotes the distribution over the L_z eigenstates. By inspection (303) becomes an equality for any eigenstate of the L_z observable. The relation was also extended to two angles φ and θ and the corresponding pair of observables L_z and L^2 (Białynicki-Birula and Madajczyk, 1985).

Another example are the number N and phase Φ observables for the harmonic oscillator. Hall (1993) shows that

$$H(N) + h(\Phi) \geq \log 2\pi \,, \tag{304}$$

where $P_N(n)$ represents the probability distribution in the number basis $\{|n\rangle\}$, and the probability distribution in the phase basis is

$$P_\Phi(\phi) := \frac{|\langle e^{i\phi}|\psi\rangle|^2}{2\pi} \quad \text{with} \quad |e^{i\phi}\rangle := \sum_n e^{in\phi}|n\rangle \,. \tag{305}$$

This can also be seen as a special case of the results in (Białynicki-Birula and Mycielski, 1975). Equation (304) becomes an equality for number states. Furthermore, Hall (1994) also extends (304) to noisy harmonic oscillators degraded by Gaussian noise.

Finally, time-energy entropic uncertainty relations for systems with discrete energy spectra were discussed in (Hall, 2008).

VI. APPLICATIONS

A. Quantum randomness

Randomness is a crucial resource for many everyday information processing tasks, ranging from online gambling to scientific simulations and cryptography. Randomness is a scarce resource since computers are designed to perform deterministic operations. Even more importantly classical physics is deterministic, meaning that every outcome of an experiment can in principle be predicted by an observer who has full knowledge of the initial state of the physical system and the operations that are performed on it. The study of pseudorandomness (Vadhan, 2012) tries to circumvent this problem.

Quantum mechanics with its inherent nondeterminism allows us to consider a stronger notion of randomness, namely randomness that is information-theoretically secure. Formally, we want to generate a random variable L that is uniformly distributed over all bit strings $\{0,1\}^\ell$ of a given length ℓ. Moreover, we want that this random variable is independent of any side information an observer might have, including information about the process that is used to calculate L and any random seeds that are used prepare L. A classical-quantum product state

$$\pi_{LE} = \frac{1}{2^\ell} \sum_{i=1}^{2^\ell} |i\rangle\langle i|_L \otimes \pi_E \tag{306}$$

describes ℓ bits of uniform randomness that is independent of its environment, or side information, E. Often, the best we can hope for is to approximate such a state. Namely, we say that ρ_{ZE} describes a state where L is δ-close to uniform on ℓ bits and independent of E if

$$\left\| \rho_{LE} - \frac{1}{2^\ell} \sum_{i=1}^{2^\ell} |i\rangle\langle i|_L \otimes \rho_E \right\|_{\mathrm{tr}} \leq \delta \,, \tag{307}$$

where $\|\cdot\|_{\mathrm{tr}}$ denotes the trace norm. This bound implies that L cannot be distinguished from a uniform and independent random variable with probability more than $\frac{1}{2}(1+\delta)$. This viewpoint is at the core of universally composable security frameworks (Canetti, 2001; Unruh, 2010), which ensure that a secret key satisfying the above property can safely be employed in any cryptographic protocol requiring a secret key.

Entropic uncertainty relations can help us since they certify that the random variables resulting from a quantum measurement are uncertain and thus contain randomness. However, in order to extract approximately uniform and independent randomness we will need an additional step, which we describe next.

1. The operational significance of conditional min-entropy

The importance of the min-entropy in cryptography is partly due to the following lemma, called the *Left-over Hashing Lemma* (Impagliazzo *et al.*, 1989; Impagliazzo and Zuckerman, 1989; McInnes, 1987). Informally, it states that there exists a family of functions $\{f_s\}_s$,

where $f_s : \mathcal{X} \to [2^\ell]$, called hash functions, such that the random variable $L = f_S(X)$, which results by applying the function f_S with S a seed chosen uniformly at random, is close to uniform and independent of S if the initial min-entropy is sufficiently large.

More formally, Renner and König (2005) and Renner (2005) show the following result for the quantum case: There exists a family $\{f_s\}_s$ as described above such that for any cq state

$$\rho_{XE} = \sum_x P_X(x)|x\rangle\langle x|_X \otimes \rho_E^x \tag{308}$$

with $H_{\min}(X|E) \geq k$, the classical-quantum-classical state ρ_{LES} after applying f_S, namely

$$\rho_{LES} = \sum_{s,x} \frac{P_X(x)}{|S|} |f_s(x)\rangle\langle f_s(x)|_L \otimes \rho_E^x \otimes |s\rangle\langle s|_S \,, \tag{309}$$

describes a state where L is δ-close to uniform on ℓ bits and independent of E and S with $\delta = 2^{\frac{1}{2}(\ell-k)}$.

The special case where the environment E is trivial has been discussed extensively in the computer science literature (Vadhan, 2012). Since hashing is a classical process, one might expect that the physical nature of the side information is irrelevant and that a classical treatment is sufficient. However, this is not true in general. For example, the output of certain extractors may be partially known if side information about their input is stored in a quantum memory, while the same output is almost uniform conditioned on any classical side information.[26]

A generalization of this result is possible by considering a variation of the min-entropy, which is called *ε-smooth* min-entropy, and denoted $H_{\min}^\varepsilon(X|E)$, for a small $\varepsilon > 0$. This is defined by maximizing the min-entropy over states that are in a ball of radius ε around the state ρ. (See Tomamichel *et al.* (2010) for a precise definition.)

The generalized Leftover Hashing Lemma (Renner, 2005) asserts that there exists a family $\{f_s\}_s$ such that for any state ρ_{XE} with $H_{\min}^\varepsilon(X|E) \geq k$, we find that $L = f_S(X)$ is $(\varepsilon + \delta)$-close to uniform and independent of E and S, with δ as defined above.

The latter result is tight in the following sense. If $L = f_S(X)$ is ε-close to uniform and independent from E and S for any family of functions $\{f_s\}_s$, then we must have $H_{\min}^{\varepsilon'}(X|E) \geq \ell$ with $\varepsilon' = \sqrt{2\varepsilon}$.

Due to this tightness result we are justified to say that the smooth min-entropy characterized (at least approximately) how much uniform randomness can be extracted from a random source X that is correlated with its environment E.

[26] See (Gavinsky *et al.*, 2009) for a concrete example and (König and Renner, 2011) for a more general discussion of this topic.

2. Certifying quantum randomness

Note that we can certify randomness, if we can somehow conclude that $H_{\min}(X|E)$ is large. In principle, all entropic uncertainty relations that involve a quantum memory are suitable for this task, whenever we can verify the terms lower bounding the entropy. Tripartite uncertainty relations are especially suitable to this task, and the security of quantum key distribution below rests on our ability to make such estimates. For example, Vallone *et al.* (2014) specialize the uncertainty relation for min- and max-entropy in (226) to assert that

$$H_{\min}(X|E)_\rho \geq \log d - H_{\max}(Z) \,, \tag{310}$$

where $\mathbb{X}$ and $\mathbb{Z}$ are mutually unbiased measurements on a d-dimensional Hilbert space. Here E is the environment of the measured system and the max-entropy $H_{\max}(Z) = H_{1/2}(Z)$ can be estimated using statistical tests, resulting in confidence about $H_{min}(X|E)$. As discussed above, the Leftover Hashing Lemma now allows to extract uniform randomness from X.

Miller and Shi (2014) derive a lower bound on an entropy difference instead of a conditional entropy. Assume that $\mathbb{X}$ and $\mathbb{Z}$ are complementary binary measurements on a qubit. Then, the following relation holds,

$$H_\alpha(XB) - H_\alpha(B) \geq q(\alpha, \delta) \quad \text{for} \quad \alpha \in (1, 2] \,, \tag{311}$$

where δ is determined by the equality

$$\operatorname{tr}\left[\langle \mathbb{Z}^0|\rho_{AB}|\mathbb{Z}^0\rangle^\alpha\right] = \delta \operatorname{tr}\left[\rho_B^\alpha\right] , \tag{312}$$

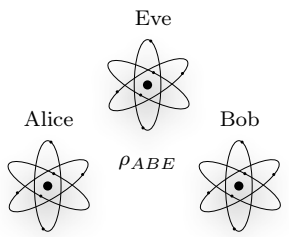
and q is a function satisfying $\lim_{\alpha\to 1} q(\alpha, \delta) = 1 - 2h(\delta)$. The authors then proceed to use this result to bound the smooth min-entropy and apply the generalized Leftover Hashing Lemma.

B. Quantum key distribution

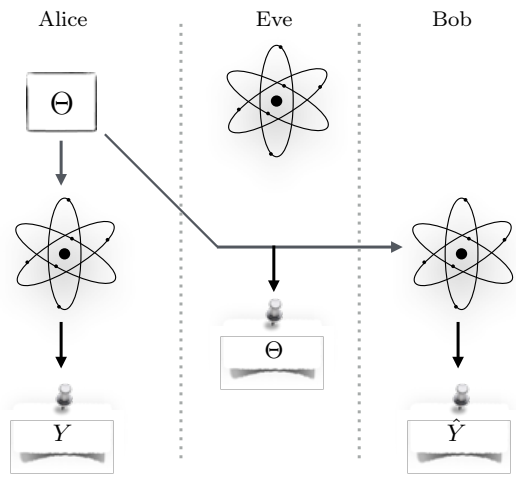
The goal of a key distribution scheme is for two honest parties to agree on a shared key by communicating over a public channel in such a way that the key is secret from any potential adversary eavesdropping on the channel. Traditionally the two honest parties trying to share a key are called Alice and Bob and the eavesdropper is called Eve. By a simple symmetry argument it is evident that key distribution is impossible if only classical information is considered: Since Eve will hear all communication from Alice to Bob, at any point in the protocol she will have at least as much information about Alice's key as Bob — in particular, if Bob knows Alice's key then so does Eve.

Quantum key distribution (QKD) was first proposed by Bennett and Brassard (1984) and Ekert (1991) to get out of this impasse.[27] Since quantum information cannot be copied or cloned (Wootters and Zurek, 1982), the

[27] We refer to (Scarani *et al.*, 2009) for a recent review.



(a) Preparation phase



(b) Measurement phase: Alice and Bob measure their quantum system in the basis indicated by Θ to recover Y and $\hat{Y}$, respectively. Eve stores Θ and keeps her quantum memory intact. The uncertainty relation is applied to the resulting state $\rho_{Y\hat{Y}E\Theta}$.

FIG. 15 Preparation and measurement phase of the QKD protocol described in Sec. VI.B.1.

above impossibility argument does no longer apply when Alice and Bob are allowed to communicate over a quantum channel. Roughly speaking, the main idea is that whenever the eavesdropper interacts with the channel (for example by performing a measurement on a particle) she will necessarily introduce noise in the quantum communication between Alice and Bob, which they can then detect and subsequently abort the protocol.

1. A simple protocol

We will focus on a truncated version of Ekert's protocol (Ekert, 1991), which proceeds as follows:

Preparation: Alice and Bob share a maximally entangled two-qubit state using the public channel. Eve can coherently interact with the channel.

Measurement: They randomly agree (using the public channel) on either the basis $\mathbb{Z} = \{|0\rangle\langle 0|, |1\rangle\langle 1|\}$ or $\mathbb{X} = \{|+\rangle\langle +|, |-\rangle\langle -|\}$, and measures their respective qubits in this basis. (These two steps are repeated many times.)

Parameter estimation: Alice announces her measurement results on a random subset of these rounds. If their measurement results agree on most rounds, they conclude that their correlations contain some secrecy and proceed to correct their errors and extract a secret key (we will not discuss this further here). If not, they abort the protocol.

2. Security criterion for QKD

To show security of QKD we thus need to show that the following two statements are mutually exclusive: a) Alice's and Bob's measurement results agree in most rounds, and b) Eve has a lot of information about Alice's or Bob's measurement outcomes.

Security of quantum key distribution against general attacks was first formally established by Mayers (1996, 2001) as well as Biham *et al.* (2000, 2006) and Shor and Preskill (2000). In all these security arguments, the complementarity or uncertainty principle is invoked in some form to argue that if Alice and Bob have large agreement on the qubits measured in one basis, then necessarily Eve's information about the bits measured in the complementary basis must be low.

In Sec. VI.B.3 we attempt to present the security argument in a concise and intuitive way, and for this purpose we adopt a notion of security—certifying that the raw key has high entropy—that has proven to be insufficient in practice. Note that our ultimate goal is to extract a secret key, and not to have a bit string with high entropy. This ultimately requires the use of different entropies and a post-processing step in the protocol to distill a secret key. A discussion of these issues follows in Sec. VI.B.4.

Entropic uncertainty relations were first used in this context by Koashi (2006), who establishes security of QKD by leveraging the Maasseen-Uffink relation (31). However, entropic uncertainty relations with quantum memory provide a more direct tool to formalize security arguments for QKD, as we will see in the following.

3. Proof of security via an entropic uncertainty relation

a. Single round. We will here broadly follow an argument outlined by Berta *et al.* (2010). First, note that during the preparation step (as described above) the eavesdropper might interfere and we can thus not know if Alice and Bob will indeed share a maximally entangled state after the preparation step is complete. However, without loss of generality we may assume that Alice (A), Bob (B), and Eve (E) share an arbitrary state ρ_{ABE} after the preparation step, where A and B are qubits and E is any quantum system held by Eve (see Fig. 15(a)).

Let Θ be a binary register in a fully mixed state that determines if the qubits are to be measured in the basis $\mathbb{X}$ or $\mathbb{Z}$ and let Y denote the output of Alice's measurement. Then we can write $H(Y|B\Theta) = \frac{1}{2}H(X|B) + \frac{1}{2}H(Z|B)$ and similarly $H(Y|E\Theta) = \frac{1}{2}H(X|E) + \frac{1}{2}H(Z|E)$. Thus, the tripartite entropic uncertainty principle with quantum memory (209) can be cast into the form

$$H(Y|E\Theta) + H(Y|B\Theta) \geq \log \frac{1}{c} = 1, \tag{313}$$

where we used that $c = \frac{1}{2}$ for the measurements $\mathbb{X}$ and $\mathbb{Z}$. The entropies are evaluated for the state $\rho_{Y\Theta BE}$ after the measurement on Alice's qubit is performed.

Next we perform Bob's measurement, which yields an estimate $\hat{Y}$ of Y. The data-processing inequality (C5) implies that $H(Y|B\Theta) \leq H(Y|\hat{Y})$, and thus we conclude that $H(Y|E\Theta) \geq 1 - H(Y|\hat{Y})$. This ensures that Eve's uncertainty—in terms of von Neumann entropy—of Alice's measurement outcome is large as long as the conditional entropy $H(Y|\hat{Y})$ is small (see Fig. 15(b)). This is a quantitative expression of the above-mentioned security criterion.[28]

Example 33. *If Alice and Bob's measurement outcomes agree with high probability, let us say with probability* $1-\delta$*, then* $H(Y|\hat{Y})$ *evaluates to* $h(\delta) = \delta \log \frac{1}{\delta} + (1-\delta)\log \frac{1}{1-\delta}$*, the binary entropy of* δ*. Hence, we find that*

$$H(Y|E\Theta) \geq 1 - h(\delta)\,, \tag{314}$$

which is positive as long as δ *is strictly less than* 50%.

b. Multiple rounds. The protocol extends over multiple rounds and we can repeat the above argument for each round individually and then attempt to add up the resulting entropies — but it is much more convenient to use a stronger uncertainty relation that describes the situation for multiple rounds directly.

For this purpose, let us model the situation after Alice and Bob have exchanged n qubits but before they measure them. This is a hypothetical situation since in the actual protocol Alice and Bob measure their qubits after every round. However, we can always imagine that Alice and Bob delay their measurement since Eve's strategy cannot depend on the timing of their measurement. After the exchange Alice and Bob each hold n qubits in systems $A^n = A_1A_2\ldots A_n$ and $B^n = B_1B_2\ldots B_n$, respectively. This is described by an arbitrary state $\rho_{A^nB^nE}$ where E is any quantum system held by the eavesdropper. Again, we model the random measurement choice using a register, a bit string $\Theta^n = (\Theta_1, \Theta_2, \ldots, \Theta_n)$ of length n in a fully mixed state, where Θ_i determines the choice of measurement on the systems indexed by i. Analogously, we store the measurement outcomes on Alice's system in a bit string $Y^n = (Y_1, Y_2, \ldots, Y_n)$ and on Bob's system in a bit string $\hat{Y}^n = (\hat{Y}_1, \hat{Y}_2, \ldots, \hat{Y}_n)$.

The crucial observation is that the tripartite uncertainty principle in (209) implies that

$$\begin{aligned} &H(X_1X_2Z_3X_4\ldots X_{n-1}Z_n|E) \\ &\quad + H(Z_1Z_2X_3Z_4\ldots Z_{n-1}X_n|B) \geq n\,, \end{aligned} \tag{315}$$

where we made sure that all n systems are measured in the opposite basis in the two terms, and used that $\log \frac{1}{c^n} = n$. A similar averaging argument as for the one round case and the data-processing inequality (C5) then reveal the bounds

$$\begin{aligned} &H(Y^n|E\Theta^n) + H(Y^n|\hat{Y}^n) \\ &\quad \geq H(Y^n|E\Theta^n) + H(Y^n|B^n\Theta^n) \geq n\,. \end{aligned} \tag{316}$$

Hence, Eve's uncertainty (in terms of von Neumann entropy) of the measurement outcome Y^n increases linearly in the number of rounds. Notably, this is true without assuming anything about the attack. In particular, the state $\rho_{A^nB^nE}$ after preparation but before the uncertainty principle is applied does not need to have any particular structure and is assumed to be arbitrary.

4. Finite size effects and min-entropy

So far we have argued that security of QKD is ensured if Eve's uncertainty of the key expressed in terms of the von Neumann entropy is large. This might be a reasonable *ad-hoc* criterion — but more operationally what we want to say is that a key is secure if it can be safely used in any other protocol, for example one-time pad encryption, that requires a secret key. This leads to the notion of composable security, which was recently reviewed by Portmann and Renner (2014) in the context of QKD. It turns out that in order to achieve composable security, it is not sufficient to consider Eve's uncertainty in terms of the von Neumann entropy (König *et al.*, 2007). Instead, it is necessary to ensure that the smooth min-entropy of the measurement results is large (Renner, 2005; Renner and König, 2005), so that we can extract a secret key, i.e. uniform randomness that is independent of the eavesdropper's memory. (Recall the discussion of randomness in Sec. VI.A.) Thus, instead of the inequality involving von Neumann entropies above (315), we want to apply a generalization of the Maassen-Uffink uncertainty relation with quantum memory (225). This leads to the following relation (Tomamichel and Renner, 2011).

$$H^{\varepsilon}_{\min}(Y^n|E\Theta^n) + H^{\varepsilon}_{\max}(Y^n|\hat{Y}^n) \geq n\,, \tag{317}$$

where $H^{\varepsilon}_{\min}$ and $H^{\varepsilon}_{\max}$ denote the smooth min- and max-entropies, variations of the min- and max-entropy (that

[28] Note that in practice we need a stronger statement, namely a bound on the min-entropy. This is discussed in Sec. VI.B.4.

we will not discuss further here). Hence, in order to ensure security it is sufficient to estimate the smooth max-entropy $H^{\varepsilon}_{\max}(Y^n|\hat{Y}^n)$. This can be done by a suitable parameter estimation procedure, as is shown in (Tomamichel *et al.*, 2012).

5. Continuous variable QKD

Quantum information processing with continuous variables (Weedbrook *et al.*, 2012) offers an interesting and practical alternative to finite-dimensional systems. Here we discuss a particular variation of the above QKD protocol where Alice and Bob measure the quadrature components of an electromagnetic field, and then extract a secret key from the correlations contained in the resulting continuous variables.

If Alice and Bob share a squeezed Gaussian state, Furrer *et al.* (2012) show that the security of such protocols can be shown rigorously using entropic uncertainty relations, including finite size effects. For this purpose, it is convenient to employ a smoothed extension of (294) as first shown by Furrer *et al.* (2011). This yields

$$H^{\varepsilon}_{\min}(Y^n|E\Theta^n) + H^{\varepsilon}_{\max}(Y^n|\hat{Y}^n) \geq n \log\left(\frac{2\pi}{\delta^2}\cdot S_0^{(1)}\left(1,\frac{\delta^2}{4}\right)^{-2}\right), \tag{318}$$

where Y_i is the outcome of the quadrature measurement in the basis (position or momentum) specified by Θ_i discretized with bin size δ. We point to (Gehring *et al.*, 2015) for an implementation.

C. Two-party cryptography

In this section we discuss applications of entropic uncertainty relations to cryptographic tasks between two mutually distrustful parties (traditionally called Alice and Bob). This setup is in contrast to quantum key distribution where Alice and Bob do trust each other and only a third party is eavesdropping. Typical tasks for two-party cryptography are bit commitment, oblivious transfer or secure identification (see Broadbent and Schaffner (2015) for a recent review).

It turns out, however, that even using quantum communication it is only possible to obtain security if we make some assumptions about the adversary (Lo, 1997; Lo and Chau, 1997; Mayers, 1997). What makes this problem harder is that unlike in QKD where Alice and Bob trust each other to check on any eavesdropping activity, here every party has to fend for himself. Nevertheless, since tasks like secure identification are of great practical importance one is willing to make such assumptions in practise.

Classically, such assumptions are typically computational assumptions: we assume a particular problem such as factoring is difficult to solve, and in addition that the adversary has limited computational resources, in particular not enough to solve the difficult problem. However, it is also possible to obtain security based on *physical* assumptions, where we will first consider assuming that the advsersies memory resources are limited. Even a limit on classical memory can lead to security Cachin and Maurer (1997) and Maurer (1992), however classical memory is typically cheap and plentiful. More significantly, however, Dziembowski and Maurer (2004) have shown that any classical protocol in which the honest players need to store n bits to execute the protocol can be broken by an adversary who is able to store more than $O(n^2)$ bits. Motivated by this unsatisfactory gap it is an evident question to ask if quantum communication can be of any help. The situation is rather different if we allow quantum communication: we can have quantum protocols (see below) that require *no* quantum memory to be executed, but that are secure as long as the adversary's quantum memory is not larger than $n - O(\log^2 n)$ qubits Dupuis *et al.* (2015) where n is the number of qubits sent during the protocol. This is essentially optimal, since any protocol that allows the adversary to store n qubits is known to be insecure Lo (1997); Lo and Chau (1997); and Mayers (1997). The asssumption of a memory limitation is known as the bounded Damgaard *et al.* (2008), or more generally, noisy-storage model Wehner *et al.* (2008), as illustrated in Fig. 16.

Security proofs in this model are intimately connected to entropic uncertainty relations. What's more, uncertainty relations of Dupuis *et al.* (2015) together with König *et al.* (2012) demonstrate that any physical assumption that limits the adversary's entanglement leads to security.

1. Weak string erasure

The relation between cryptographic security and entropic uncertainty relations can easily be understood by looking at the very simple cryptographic building block known as weak string erasure (WSE) (König *et al.*, 2012). Weak string erasure is universal for two-party secure computation, in the sense that any other protocol can be obtained by repeated executions of weak string erasure, following by additional quantum or classical communication (Kilian, 1988). Importantly, the storage assumption only needs to hold during some time Δt during the execution of weak string erasure.

Weak string erasure generates the following outputs if both Alice and Bob are honest: Alice obtains an n-bit string K^n, and Bob obtains a random subset $I \subseteq [n]$, and the bits $K_I \subseteq K^n$ of K^n as indexed by the subset I. In addtition, the following demands are made for security. If Bob is honest, then Alice does not know anything about I. In turn, if Alice is honest, then Bob

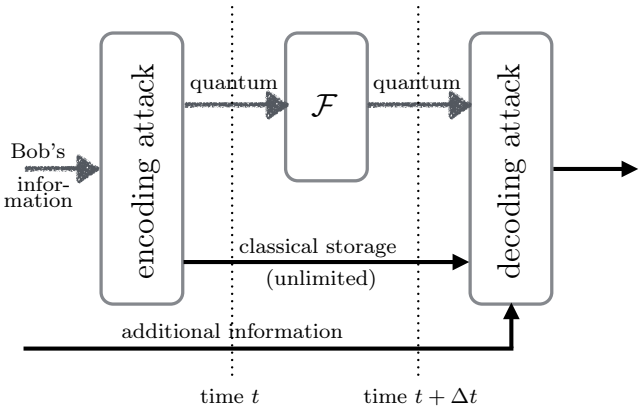


FIG. 16 The noisy-storage model König *et al.* (2012) and Wehner *et al.* (2008) assumes that during waiting times Δt in the protocol, the adversary can only keep quantum information in an imperfect and limited storage device described by a quantum channel, $\mathcal{F}$. This is the only restriction and the adversary is otherwise abitrarily powerful. In particular, he can first store all incoming qubits, and has a quantum computer to encode them into an abitrary quantum error-correcting code to protect them against the noise of the channel $\mathcal{F}$. He can also keep an unlimited amount of classical memory, and perform any operation instantaneously.

should not too much about K^n (except for K_I). More precisely, Bob should not be able to guess K^n too well, that is with (140),

$$H_{\min}(K^n|B) \geq \lambda \cdot n \quad \text{for some} \quad \lambda \in [0,1]\,, \tag{319}$$

where B denotes all of Bob's knowledge. We refer to König *et al.* (2012) for a more detailed definition. A simple protocol for implementing weak string erasure is as follows:

1. Alice prepares a random n bit string K^n, encodes each bit K_i in one of the BB84 Bennett and Brassard (1984) bases $\Theta \in \{\sigma_{\mathbb{X}}, \sigma_{\mathbb{Z}}\}$ at random, and sends these n qubits to Bob.

2. Bob measures the n qubits in randomly chosen bases $\Theta' \in \{\sigma_{\mathbb{X}}, \sigma_{\mathbb{Z}}\}$.

3. After the waiting time Δt, Alice sends the classical n bit string Θ^n to Bob and outputs K^n.

4. Bob computes $I = \{i : \theta_i = \theta'_i\}$ and outputs I and K_I.

Note that if both parties are honest, then the protocol is correct in the sense that Alice outputs K^n and Bob I with $K_I \subseteq K^n$. Moreover, when Alice is dishonest, it is intuitively obvious that she is unable to gain any information about the index set I (even with an arbitrary quantum memory), since she never receives any information from Bob during the protocol. A precise argument for this can be found in, e.g., König *et al.* (2012). On the other hand, note that a dishonest Bob with a quantum memory can easily cheat by just keeping the n qubits he gets from Alice and wait until he receives the n bit string Θ^n from Alice as well. Namely, he can then measure the n qubits in the same basis Θ^n as Alice and get the full n bit string K^n (that is: $H_{\min}(K^n|B\Theta^n) = 0$). However, if Bob only has a limited quantum memory, then he could not keep a perfect copy of the n qubits he gets from Alice.

The security analysis is linked immediately to a guessing game, whenever we consider a purified version of the protocol in which Alice does not prepare BB84 states herself, but instead makes EPR pairs $|\Phi_{AB}\rangle = (|0\rangle_A|0\rangle_B + |1\rangle_A|1\rangle_B)/\sqrt{2}$ and sends B to Bob, while measuring A in a randomly chosen BB84 basis. In the analysis, one can indeed give even more power to Bob, in which we imagine that he prepares a state ρ_{AB} in each round of the protocol and Alice measures A in randomly chosen BB84 basis. Alice then sends him the basis choice. Recall that $H_{\min}(K^n|B\Theta^n) = -\log p_{\text{guess}}(K^n|B\Theta^n)$, that is, the min-entropy security guarantee that WSE demands is then precisely related to Bob's ability to win the guessing game Ballester *et al.* (2008). The storage assumption then translates into one particular example of how the entanglement in ρ_{AB} is limited, putting a limit on $H_{\min}(A^n|B)$ of the states that Bob can prepare.

2. Bounded-storage model

To illustrate further how a bound on entropic uncertainty leads to security, let us first consider a special case of the noisy-storage model, also known as the bounded-storage model. Here, the channel $\mathcal{F} = \mathcal{I}_2^{\otimes q}$ in Fig. 16 is just the identity on q qubits. This *bounded-storage model* was introduced and first studied in Damgaard *et al.* (2007, 2008); and Schaffner (2007).

While much more refined bounds are known Dupuis *et al.* (2015), let us first explain how already entropic uncertainty relations for a classical memory system can be used to obtain weak security statements in this setting. To this end, we first differentiate Bob's knowledge into $B = QM\Theta^n$, where Q denotes the q qubits of quantum memory, M denotes (unbounded) classical information, and Θ^n is the n bit basis information string Alice sent to Bob. Since the conditional min-entropy obeys a chain rule Renner (2005), we can separate the quantum memory as

$$H_{\min}(K^n|B) = H_{\min}(K^n|QM\Theta^n) \tag{320}$$
$$\geq H_{\min}(K^n|M\Theta^n) - q\,. \tag{321}$$

Analyzing $H_{\min}(K^n|M\Theta^n)$ is then directly determined by Bob's ability to win the guessing game, in which he only has classical information M. Using the min-entropy uncertainty relation (104) for the n qubit BB84 measurements (with an extension to classical side information M

as sketched in Sec. IV.C), we get

$$H_{\min}(K^n|M\Theta^n) \geq -n \cdot \log\left(\frac{1}{2} + \frac{1}{2\sqrt{2}}\right) . \tag{322}$$

Hence, we find a non-trivial lower bound

$$H_{\min}(K^n|B) > 0 \quad \text{as long as} \quad q \lesssim n \cdot 0.22 . \tag{323}$$

This security analysis can be refined and improving on Damgaard *et al.* (2007), Ng *et al.* (2012) make use of the following stronger smooth min-entropy uncertainty relation which is based on (105),

$$\begin{aligned} &H^{\varepsilon}_{\min}(K^n|M\Theta^n) \\ &\geq n \cdot \sup_{s\in(0,1]} \left(\frac{1}{s}\big(1 + s - \log\left(1 + 2^s\right)\big) - \frac{1}{sn}\log\frac{2}{\varepsilon^2}\right) . \end{aligned} \tag{324}$$

One can use this uncertainty relation together with the more refined analysis of König *et al.* (2012) instead of (321), to obtain perfect security ($\lambda \to 1$) against quantum memory of size

$$q \leq \frac{n}{2} . \tag{325}$$

for $n \to \infty$. Ultimately, Dupuis *et al.* (2015) show by deriving strong entropic uncertainty relations that the protocol from Sec. VI.C.1 implements a WSE scheme against q qubits of quantum memory for

$$\lambda = \frac{1}{2}\left(\gamma_{BB84}\left(-\frac{q}{n}\right) - \frac{1}{n}\right) , \tag{326}$$

where the function $\gamma_{BB84}(\cdot)$ is as in (205). Asymptotically ($n \to \infty$), this provides perfect security ($\lambda \to 1$) against quantum memories of size

$$q \leq n - O\left(\log^2 n\right) . \tag{327}$$

This is basically optimal, since no protocol can be secure if $q = n$. Finally, we mention that alternatively we could also use a six state encoding $\{\sigma_{\mathbb{X}}, \sigma_{\mathbb{Y}}, \sigma_{\mathbb{Z}}\}$ for the weak string erasure protocol described in Sec. VI.C.1. We refer to Dupuis *et al.* (2015); Mandayam and Wehner (2011); and Ng *et al.* (2012) for a security analysis.

3. Noisy-storage model

Let us know consider the most general case of arbitrary storage devices $\mathcal{F}$ in Fig. 16 (Wehner *et al.*, 2008). This model is motivated by the fact that counting qubits is generally a significant over estimate of the storage capabilities of a quantum memory, and indeed for example for continous variable systems there is no dimension bound to which to apply the bounded-storage analysis. The first general security analysis was given by König *et al.* (2012), which was then refined significantly by Berta *et al.* (2013, 2014b), leading to the asymptotically tight security analysis by Dupuis *et al.* (2015). Here, one can not just use the chain rule to separate the quantum memory as in (320) – (321). Such a separation is only possible when relating the security to the classical capacity of the storage channel $\mathcal{F}$ (König *et al.*, 2012). Instead, we have to apply a min-entropy uncertainty relation with quantum memory to directly lower bound

$$H_{\min}(K^n|B) = H_{\min}(K^n|QM\Theta^n) . \tag{328}$$

We use a variant of the relation (204) for the n qubit BB84 measurements to bound (Dupuis *et al.*, 2015),

$$\begin{aligned} H^{\varepsilon}_{\min}(K^n|QM\Theta^n) \geq n \cdot \gamma_{BB84}&\left(\frac{H_{\min}(A^n|QM)}{n}\right) \\ &- 1 - \log\left(\frac{2}{\varepsilon^2}\right) , \end{aligned} \tag{329}$$

where the function $\gamma_{BB84}(\cdot)$ is as in (205). In order to get an idea how to lower bound the right-hand side of (329) under a noisy quantum memory Q assumption, recall that $H_{\min}(A^n|QM)$ is a measure of entanglement between A^n and $B = QM$. In particular, one can relate this amount of entanglement to Bob's ability to store the n EPR pairs that Alice sends in the purified version of the protocol, that is, the quantum capacity of the storage channel $\mathcal{F}$. If $\mathcal{F}$ cannot preserve said entanglement, then $H_{\min}(K^n|QM\Theta^n)$ in (329) will be lower bounded non-trivially leading to a secure WSE scheme for some trade-off between the security parameter λ from (319), the number n of qubits sent, and the noisiness of the quantum memory Q. We refer to Dupuis *et al.* (2015) for any details.

Again we could also use a six state encoding $\{\sigma_{\mathbb{X}}, \sigma_{\mathbb{Y}}, \sigma_{\mathbb{Z}}\}$ for the weak string erasure protocol described in Sec. VI.C.1. We refer to Berta *et al.* (2014b) and Dupuis *et al.* (2015) for a security analysis.

4. Uncertainty in other protocols

Entropic uncertainty relations feature in many other quantum cryptographic protocols. The entropic relation for channels (245) was used in Buhrman *et al.* (2008) to obtain cheat-sensitivity for a quantum string commitment protocol. The same relations as relevant for the noisy-storage model above, have also been used to prove security in the isolated qubit model Liu (2014, 2015). In this model, the adversary is given a quantum memory of potentially long-lived qubits, but they are isolated in the sense that he is unable to perform coherent operations on many qubits simultaneously. In particular, the uncertainty relation of Damgaard *et al.* (2007) was used in Liu (2014) to obtain security. It would possible to use the relation (105) from Ng *et al.* (2012) to obtain improved security parameters. Furthermore, tripartite Tomamichel

et al. (2013) and bipartite Ribeiro and Grosshans (2015) uncertainty relations have been used to ensure the security of position-based cryptography. Finally, in relativistic cryptography, security of two-party protocols is possible under the assumptions that each player is split into several non-communicating agents, and tripartite uncertainty relations have been used to establish security in this setting (Kaniewski *et al.*, 2013).

D. Entanglement witnessing

Entanglement is a central resource in quantum information processing, hence methods for detecting entanglement are crucial for quantum information technologies. Entanglement witnessing refers to the process of verifying that a source is producing entangled particles. Entangled states are defined as those states that are non-separable, i.e., they cannot be written as a convex combination of product states. A common theme in entanglement witnessing is to first prove a mathematical identity that all separable states must satisfy; let us refer to such an identity as an entanglement witness. Experimentally demonstrating that one's source violates this identity will then guarantee that the source produces entangled particles.

Entanglement witnessing is a well-developed field (e.g., see the review article by Gühne and Tóth (2009) and Horodecki *et al.* (2009)), and their are many types of entanglement witnesses. Here we focus on entanglement witnesses that follow from entropic uncertainty relations.

In what follows, we restrict the discussion to bipartite entanglement. We note that entanglement witnessing typically occurs in the distant-laboratories paradigm, where two parties - Alice and Bob - can each perform local measurements on their respective systems, but neither party can perform a global measurement on the bipartite system.

For introductory purposes, let us mention a simple, well-known bipartite entanglement witness for two qubits. Although it is non-entropic, it is based on complementary observables, and so it can be directly compared to the entropic witnesses discussed below. Namely, consider the operator

$$E_{XZ} := E_X + E_Z, \text{ where} \tag{330}$$
$$E_X := |+\rangle\langle+| \otimes |-\rangle\langle-| \;+\; |-\rangle\langle-| \otimes |+\rangle\langle+|, \tag{331}$$
$$E_Z := |0\rangle\langle 0| \otimes |1\rangle\langle 1| + |1\rangle\langle 1| \otimes |0\rangle\langle 0|\,. \tag{332}$$

Note that E_X and E_Z are "error operators" in that they project onto the subspaces where Alice's and Bob's measurement outcomes are different. For a maximally entangled state of the form $|\psi\rangle = (|00\rangle + |11\rangle)/\sqrt{2}$, there is no probability for error in either basis, so we have $\langle\psi|E_{XZ}|\psi\rangle = 0$. On the other hand, for any separable state ρ_{AB}, we have that (e.g., see Namiki and Tokunaga (2012))

$$\mathrm{tr}[\rho_{AB}E_{XZ}] \geq \frac{1}{2}\,. \tag{333}$$

Hence if $\langle E_X\rangle + \langle E_X\rangle < 1/2$, where $\langle O\rangle := \mathrm{tr}[O\rho_{AB}]$, then ρ_{AB} is entangled. This witness is depicted as the solid line in Fig. 17.

1. Shannon entropic witness

Consider the bipartite uncertainty relation with quantum memory in (167). Berta *et al.* (2010) discussed how this can be used for entanglement witnessing, and this approach was experimentally implemented by Li *et al.* (2011) and Prevedel *et al.* (2011). Specifically, from (167), one finds that all separable states satisfy

$$H(X_A|X_B) + H(Z_A|Z_B) \geq \log\frac{1}{c}\,, \tag{334}$$

where the c parameter refers to Alice's observables, whereas Bob's observables $\mathbb{X}_B$ and $\mathbb{Z}_B$ are arbitrary. One can see this by noting that $H(A|B) \geq 0$ for any separable state, and furthermore that measuring Bob's system in some basis $\mathbb{X}_B$ cannot reduce his uncertainty about Alice's measurement, i.e., $H(X_A|X_B) \geq H(X_A|B)$.

One can use (334) for entanglement witnessing, using a protocol where Alice and Bob have many copies of ρ_{AB} and they each measure either their $\mathbb{X}$ or $\mathbb{Z}$ observable on each copy. The quantities $H(X_A|X_B)$ and $H(Z_A|Z_B)$ can then be calculated from their joint probability distributions $\Pr(X_A = x_A, X_B = x_B)$ and $\Pr(Z_A = z_A, Z_B = z_B)$, and if (334) is violated, then ρ_{AB} must be entangled.

Fig. 17 depicts this entanglement witness (long-dashed curve) for the case of qubits and mutually unbiased bases. A comparison of this curve to the black line shows that''(333) detects more entangled states than (334). However, the "quality" of entanglement that (334) detects is higher. This is because (334) holds for all non-distillable states, i.e., states from which Alice and Bob cannot distill any EPR (maximally-entangled) states using local operations and classical communication (LOCC) (see, e.g., (Horodecki *et al.*, 2009) for a discussion of LOCC). In this sense, (334) detects *distillable* entanglement whereas (333) detects all forms of entanglement.

One can make this quantitative using a result by Devetak and Winter (2005) that minus the conditional entropy lower bounds the distillable entanglement E_D, i.e., the optimal asymptotic rate for distilling EPR states using LOCC:

$$E_D \geq -H(A|B)\,. \tag{335}$$

Combining this with (167) gives

$$E_D \geq \log\frac{1}{c} - H(X_A|X_B) - H(Z_A|Z_B)\,. \tag{336}$$

This reveals an advantage of the entropic uncertainty approach to entanglement witnessing, namely, that it can give quantitative lower bounds, in contrast to witnesses like that in (333) that only answer a "yes or no" question.

Another advantage of the entropic uncertainty approach is that it requires no structure on Bob's side. While (333) requires both Alice's and Bob's measurements to be mutually unbiased, the entropic uncertainty approach allows for arbitrary measurements on Bob's system, and hence is more flexible.

2. Other entropic witnesses

Other bipartite quantum memory uncertainty relations (see Sec. IV) lead to similar entanglement witnesses. For example, Berta *et al.* (2014a) discuss how the uncertainty relation in (188) allows for entanglement witnessing using a set of n MUBs on Alice's system (more precisely, a subset of size n of MUBs chosen from a set of d_A+1 MUBs, where d_A is a prime power and $2 \leq n \leq d_A+1$). Consider such a set $\{\mathbb{X}_j\}$ of n MUBs on Alice's system, and consider a set of n arbitrary POVMs $\{\mathbb{Y}_j\}$ on Bob's system. Berta *et al.* (2014a) show that all separable states must satisfy

$$\sum_{j=1}^{n} 2^{-H_2(X_j|Y_j)} \leq 1 + \frac{n-1}{d_A}\,. \tag{337}$$

Fig. 17 compares this entanglement witness (short-dashed curve) to the previous discussed ones, in the qubit case with $n=2$. Notice that (337) detects more entangled states than (334), but not as much as (333).

Similar to the Shannon entropy case in (336), the uncertainty relation (188) actually allows one to give a quantitative lower bound on an entanglement-like measure. Namely, (188) allows one to lower bound $-H_{\text{coll}}(A|B)$.

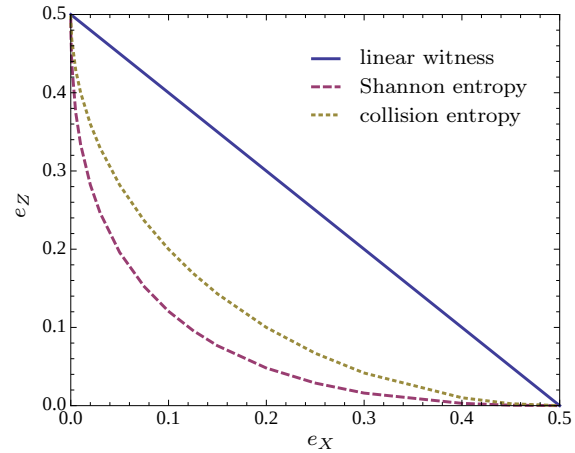


FIG. 17 Entanglement witnessing for a bipartite two-qubit state using mutually unbiased observables. Suppose Alice and Bob observe $\Pr[X_A=X_B=0]=\Pr[X_A=X_B=1]=\frac{1-e_X}{2}$ and $\Pr[X_A=0,X_B=1]=\Pr[X_A=1,X_B=0]=\frac{e_X}{2}$ and analogously for Z and e_Z. The region below the curve in the plot indicates the region for which one can guarantee entanglement for the respective witnesses.

E. EPR steering inequalities

Steering is a phenomenon for bipartite quantum systems that is related to (although not precisely the same as) entanglement. Like the previous subsection, we consider the distant-laboratories paradigm involving two parties, Alice and Bob, where Alice (Bob) has access to system A (B). Steering corresponds to one party's (say Alice's) measurement choice giving rise to different ensembles of states on the other party's (Bob's) system. Not all quantum states exhibit steering, e.g., separable states are non-steerable. At the other extreme, all states that violate a Bell inequality are steerable. While Bell inequalities are derived for states that admit a local hidden variable (LHV) model, Wiseman *et al.* (2007) formalized the notion of steerability as those states ρ_{AB} that do not admit a local hidden state (LHS) model. An LHS model is a model where, say, system B has a local quantum state that is classically correlated to arbitrary observables on system A. This formalization has led researchers to derive steering inequalities (Cavalcanti *et al.*, 2009), in analogy to Bell inequalities.

Schneeloch *et al.* (2013) and Walborn *et al.* (2011) show how entropic uncertainty relations can be used to derive steering inequalities. The idea is that if B has a local hidden state, then its measurement probabilities must obey a single system uncertainty relation, even if they are conditioned on the measurement outcomes on A. More precisely, an LHS model implies that the joint probability distribution for discrete observables $\mathbb{X}_A$ on A and $\mathbb{X}_B$ on B has the form

$$P(\mathbb{X}_A,\mathbb{X}_B) = \sum_{\lambda} P(\Lambda=\lambda)P(\mathbb{X}_A|\Lambda=\lambda)P_Q(\mathbb{X}_B|\Lambda=\lambda)\,. \tag{338}$$

Here Λ is the hidden variable that determines Bob's local state, λ is a particular value that this variable may take, and the subscript Q on $P_Q(\mathbb{X}_B|\Lambda=\lambda)$ emphasizes that the probability distribution arises from a single quantum state. Next, we have that

$$H(X_B|X_A) \geq H(X_B|X_A,\Lambda) \tag{339}$$

$$= \sum_{\lambda} P(\Lambda=\lambda)H(X_B|X_A,\Lambda=\lambda) \tag{340}$$

$$= \sum_{\lambda} P(\Lambda=\lambda)H(X_B|\Lambda=\lambda)\,, \tag{341}$$

where the notation $H(X_B|X_A,\Lambda=\lambda)$ should be read as the entropy of X_B conditioned on X_A and conditioned

on the event that $\Lambda = \lambda$. Hence for two observables $\mathbb{X}_B$ and $\mathbb{Z}_B$ on B, and some other observables $\mathbb{X}_A$ and $\mathbb{Z}_A$ on A, we have

$$H(X_B|X_A) + H(Z_B|Z_A) \geq \sum_\lambda P(\Lambda = \lambda)[H(X_B|\Lambda = \lambda) + H(Z_B|\Lambda = \lambda)] \quad (342)$$

Combining this with, say, Maassen-Uffink's uncertainty relation (31) gives the following steering inequality (Schneeloch *et al.*, 2013),

$$H(X_B|X_A) + H(Z_B|Z_A) \geq \log \frac{1}{c}\,, \quad (343)$$

where c refers to Bob's observables. Any state ρ_{AB} that admits an LHS model must satisfy (343), hence an experimental violation of (343) would constitute a demonstration of steering. Similar steering inequalities can be derived for continuous variables (Walborn *et al.*, 2011).

F. Wave-particle duality

Wave-particle duality is the fundamental concept that a single quantum system can exhibit either wave behavior or particle behavior, but the two behaviors compete, in the sense that one cannot design an interferometer that can simultaneously show both behaviors. This idea was qualitatively discussed, e.g., by Feynman, and was subsequently put on quantitative grounds by Wootters and Zurek (1979) and Jaeger *et al.* (1995) as well as Englert (1996) and Englert and Bergou (2000), who proved inequalities known as wave-particle duality relations (WPDRs). Many such relations consider the Mach-Zehnder interferometer for single photons, shown in Fig. 18. In this case, particle behavior is associated with knowing the path that the photon travels through the interferometer, while wave behavior is associated with seeing oscillations in the probability to detect the photon in a given output mode as one varies the relative phase ϕ between the two interferometer arms. Denoting the which-path observable as $\mathbb{Z} = \{|0\rangle\langle 0|, |1\rangle\langle 1|\}$, then particle behavior can be quantified by the path predictability $\mathcal{P} = 2p_{\text{guess}}(Z) - 1$, which is related to the probability of guessing the path correctly $p_{\text{guess}}(Z)$. The wave behavior is quantified by the fringe visibility

$$\mathcal{V} = \frac{p^0_{\max} - p^0_{\min}}{p^0_{\max} + p^0_{\min}}\,, \quad (344)$$

where p_0 is the probability for the photon to be detected by detector D_0 (see Fig. 18), $p^0_{\max} := \max_\phi p_0$, and $p^0_{\min} := \min_\phi p_0$. Wootters and Zurek (1979) prove that

$$\mathcal{P}^2 + \mathcal{V}^2 \leq 1\,, \quad (345)$$

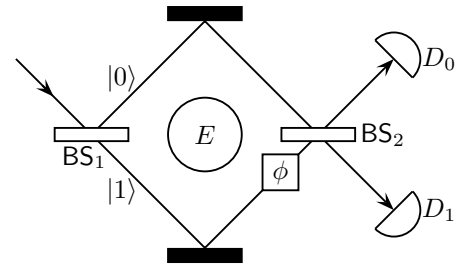


FIG. 18 Mach-Zehnder interferometer for single photons. A photon impinges on a beam splitter, after which we label the two possible paths by the $\mathbb{Z}$ basis states $|0\rangle, |1\rangle$. The photon may interact with some environment E inside the interferometer. Then a phase ϕ is applied to the lower path, and the two paths are recombined on a second beam splitter. Finally the photon is detected at either D_0 or D_1.

which implies $\mathcal{V} = 0$ when $\mathcal{P} = 1$ (full particle behavior means no wave behavior) and vice-versa.

More generally, suppose the photon may interact with some environment system E inside the interferometer. Measuring E might reveal, e.g., some information about which path the photon took, so it is natural to consider the path distinguishability

$$\mathcal{D} = 2p_{\text{guess}}(Z|E) - 1\,. \quad (346)$$

Englert (1996) and Jaeger *et al.* (1995) prove a stronger version of (345), namely

$$\mathcal{D}^2 + \mathcal{V}^2 \leq 1\,. \quad (347)$$

WPDRs such as (345) and (347) have often been thought to be conceptually different from uncertainty relations, although this has been debated.

It turns out that (345) and (347) are actually entropic uncertainty relations in disguise (Coles *et al.*, 2014). In particular, they correspond to the uncertainty relation for the min- and max-entropies, see (226), applied to complementary qubit observables. Namely, (345) is equivalent to the uncertainty relation,

$$H_{\min}(Z) + \min_{W \in XY} H_{\max}(W) \geq 1\,, \quad (348)$$

where the $\min_{W\in XY}$ corresponds to minimizing over all observables in the xy plane of the Bloch sphere. Likewise (347) is equivalent to the uncertainty relation

$$H_{\min}(Z|E) + \min_{W \in XY} H_{\max}(W) \geq 1\,. \quad (349)$$

This unifies the wave-particle duality principle with the entropic uncertainty principle, showing that the former is a special case of the latter.

G. Quantum metrology

Quantum metrology deals with the physical limits on the accuracy of measurements (Giovannetti *et al.*, 2011).

The uncertainty principle plays an important role in establishing such physical limits. Typically in quantum metrology one is interested in estimating an optical phase, e.g., the phase shift in an interferometer (as in Fig. 18). Hence uncertainty relations involving the phase observable have application here. Recall that we briefly discussed an entropic uncertainty relation for the number and phase observables in Sec. V.F, specifically in Eq. (304). While quantum metrology is a broad field (see, e.g., Giovannetti *et al.* (2011) for a review), we mention here a few works that exploit entropic uncertainty relations.

The Heisenberg limit is a well-known limit in quantum metrology stating that the uncertainty in the phase estimation scales as $1/\langle N\rangle$. Here, $\langle N\rangle$ is the mean photon number of the light that is used to probe the phase. Hall *et al.* (2012) note that the Heisenberg limit is heuristic, and they put it on rigorous footing by proving the following bound,

$$\delta\hat{\Phi} \geq k/\langle N+1\rangle\,, \tag{350}$$

where $\delta\hat{\Phi}$ is the root-mean-square deviation of the phase estimate $\hat{\Phi}$ from the actual phase Φ, and $k := \sqrt{2\pi/e^2}$. To prove (350), Hall *et al.* (2012) define the random variable $\Theta := \hat{\Phi} - \Phi$ and apply the entropic uncertainty relation in (304), giving

$$H(N) + h(\Theta) \geq \log 2\pi\,. \tag{351}$$

Then they combine (351) with some simple identities that relate $h(\Theta)$ to $\delta\hat{\Phi}$ and $H(N)$ to $\langle N+1\rangle$, to obtain (350).

Hall and Wiseman (2012) consider a more general scenario where one may have some prior information about the phase, and they likewise use the entropic uncertainty relation to obtain a rigorous statement of the Heisenberg limit.

H. Other applications in quantum information theory

Recent efforts to understand the classical-quantum boundary, in the context of both physics and information-processing, have led to quantitative measures of "quantumness" like coherence and discord, which are discussed in Sec. VI.H.1 and VI.H.2, respectively. We further discuss information locking in Sec. VI.H.3 and touch on quantum coding in Sec. VI.H.4.

1. Coherence

Baumgratz *et al.* (2014) introduced a framework for quantifying coherence, which is a measure of quantumness that satisfies some properties, such as not increasing under incoherent operations. There are a variety of coherence measures, but one in particular has an operational meaning in terms of the number of distillable maximally coherent states (Winter and Yang, 2015), namely the relative entropy of coherence

$$\Phi(\mathbb{Z},\rho) := D\left(\rho \middle\| \sum_z |\mathbb{Z}^z\rangle\langle\mathbb{Z}^z|\rho|\mathbb{Z}^z\rangle\langle\mathbb{Z}^z|\right)\,. \tag{352}$$

Note that the coherence is a function of the state ρ as well as the orthonormal basis $\mathbb{Z} = \{|\mathbb{Z}^z\rangle\langle\mathbb{Z}^z|\}$.

The following connection between coherence and entropic uncertainty was established in (Coles, 2012; Coles *et al.*, 2011). Let ρ_S be any state for system S, let $\mathbb{Z}$ be a projective observable on S, then

$$\Phi(\mathbb{Z},\rho_S) = H(Z|E)\,, \tag{353}$$

where E is a purifying system for ρ_S. This states that the relative entropy of coherence for a projective observable is equivalent to the uncertainty of that observable given the purifying system, or in other words, given access to the environment E.

The right-hand-side of (353) quantifies uncertainty in the presence of quantum memory, and uncertainty relations for such measures have been discussed in Sec. IV. Hence one can reinterpret such uncertainty relations, e.g., (167), as lower bounds on the *coherence* of ρ_S for different observables.

This idea was discussed by Korzekwa *et al.* (2014b), although they focused more on the perspective of Luo (2005) of separating total uncertainty into a "classical" and "quantum" part. In particular, for a rank-one projective observable $\mathbb{Z} = \{|\mathbb{Z}^z\rangle\langle\mathbb{Z}^z|\}$ and a quantum state ρ, they defined the classical uncertainty as the entropy of the state, $C(\mathbb{Z},\rho) := H(\rho)$, and the quantum uncertainty as the relative entropy of coherence,

$$Q(\mathbb{Z},\rho) := D\left(\rho \middle\| \sum_z |\mathbb{Z}^z\rangle\langle\mathbb{Z}^z|\rho|\mathbb{Z}^z\rangle\langle\mathbb{Z}^z|\right)\,. \tag{354}$$

It is straightforward to show that overall uncertainty is the sum of the classical and quantum parts

$$H(Z) = Q(\mathbb{Z},\rho) + C(\mathbb{Z},\rho)\,. \tag{355}$$

Korzekwa *et al.* (2014b) derive several uncertainty relations for the quantum uncertainty $Q(\mathbb{Z},\rho)$. However, using (353), one can reinterpret their relations as entropic uncertainty relations in the presence of quantum memory. In particular, their uncertainty relations follow directly from combining (167) with (353).

2. Discord

Ollivier and Zurek (2001) quantify quantum correlations by discord, which is defined as

$$D(B|A) := I(A:B) - J(B|A)\,, \tag{356}$$

which is the difference between the quantum mutual information $I(A:B)$ and the classical correlations,

$$J(B|A) := \max_{\mathbb{X}} I(X:B)\,, \tag{357}$$

where the optimization is over all POVMs $\mathbb{X}$ acting on system A. In Sec. IV.E, Ex. 25, we discussed how discord quantifies the gap between conditioning on classical versus quantum memory. Another connection to discord is the following. In an effort to strengthen the uncertainty relation with quantum memory in (167), (Pati *et al.*, 2012) introduced an additional term on the right-hand-side that depends on the discord of the state ρ_{AB}. Namely they proved the bound

$$H(X|B) + H(Z|B) \geq \log\frac{1}{c} + H(A|B) + \max\Big\{0, D(B|A) - J(B|A)\Big\}\,. \tag{358}$$

Clearly this strengthens the bound in (167) for states ρ_{AB} whose discord exceeds their classical correlations: $D(B|A) > J(B|A)$. Indeed, Pati *et al.* (2012) showed that this is true for Werner states, and furthermore that their bound (358) is perfectly tight for Werner states.

In turn, this result was used by Hu and Fan (2013) to obtain a strong upper bound on discord. That is, the uncertainty relation (358) allows one to bound the discord by

$$D(B|A) \leq \frac{1}{2}\Big(I(A:B) + \delta_T\Big)\,, \tag{359}$$

where

$$\delta_T := H(X|B) + H(Z|B) - \log\frac{1}{c} - H(A|B)\,. \tag{360}$$

Here δ_T is the gap between the left and right hand sides in the uncertainty relation (167).

3. Locking of classical correlations

One operational way of understanding entropic uncertainty relations is in terms of information locking (DiVincenzo *et al.*, 2004). In the following we present a cryptographic view on information locking as discussed by Fawzi *et al.* (2011).

A locking scheme is a protocol that encodes a classical message into a quantum state using a classical key of size much smaller than the message. The goal is that without knowing the key the message is locked in the quantum state such that any possible measurement only reveals a negligible amount of information about the message. Furthermore, knowing the key it is possible to unlock and completely recover the message. The connection of information locking to entropic uncertainty is best presented by means of a simple example based on the Maassen-Uffink bound for the n qubit BB84 measurements (103),

$$H(K^n|\Theta^n) \geq n\cdot\frac{1}{2} \;\text{ with }\; \Theta^n \in \{\theta_1,\ldots,\theta_{2^n}\}\,. \tag{361}$$

In order to encode a uniformly random n bit string X we choose at random a n qubit BB84 basis θ_i (the key) and encode the message in this basis. Based on (361), DiVincenzo *et al.* (2004) show that for any measurement on this quantum state the mutual information between the outcome of that measurement and the original classical message X is at most $n/2$. That is, $n/2$ bits are locked in the quantum state and are not accessible without knowing the basis choice (the key). This is remarkable because any non-trivial purely classical encryption of a n bit string message requires a key of size at least $n-1$. Of course, this then raises the question about the optimal trade-off between the number of lockable bits and the key size. For that purpose Fawzi *et al.* (2011) make use of the uncertainty relation (102),

$$H(K|\Theta) \geq n\cdot(1-2\varepsilon) - h(\varepsilon) \;\text{ with }\; \Theta = \{\theta_1,\ldots,\theta_L\}\,. \tag{362}$$

Based on this they show that a key size of $L = O(\log(n/\varepsilon))$ allows for locking an n bit string up to a mutual information (accessible information) smaller than $\varepsilon > 0$. State-of-the-art results use stronger definitions for information locking in terms of the trace norm instead of the mutual information and are based on so-called metric uncertainty relations (Dupuis *et al.*, 2013; Fawzi *et al.*, 2011).[29] Finally, we mention that Guha *et al.* (2014) initiated the study of the information locking capacity of quantum channels, which is also intimately related to uncertainty.

4. Quantum Shannon theory

The original partial results and conjectures for entropic uncertainty relations with quantum memory by Christandl and Winter (2005) and Renes and Boileau (2009) were inspired by applications in quantum Shannon theory (Renes and Boileau, 2008). More recently, entropic uncertainty relations and in particular their equality conditions have been used to analyze the performance of quantum Polar codes, e.g. in (Renes *et al.*, 2015; Renes and Wilde, 2014).

[29] We emphasize that the security definitions for information locking are not composable (see, e.g., (Renner, 2005) for a discussion).

VII. MISCELLANEOUS TOPICS

A. Tsallis and other entropy functions

From a mathematical perspective it is interesting to consider uncertainty relations for various generalizations of the Shannon entropy. While the Rényi entropies were discussed above, the Tsallis entropies are another family of interest. The Tsallis entropy of order α is defined as

$$H_\alpha^T(X) := \left(\frac{\log e}{1-\alpha}\right)\left(\sum_x P_X(x)^\alpha - 1\right) \quad \text{for} \quad \alpha \in (0,1)\cup(1,\infty)\,, \tag{363}$$

and as the corresponding limit for $\alpha \in \{0,1,\infty\}$. Similar to the Rényi entropies, the $\alpha = 1$ Tsallis entropy corresponds to the Shannon entropy. Note that for $x \approx 1$ we have $\log x \approx \log e \cdot (x-1)$, so when $\sum_x P_X(x)^\alpha \approx 1$ the Tsallis entropy approximates the Rényi entropy.

Rastegin has studied uncertainty relations in terms of the Tsallis entropy. For example, Rastegin (2012) proved the following uncertainty relation for Tsallis entropies, for a set of three MUBs $\{\mathbb{X},\mathbb{Y},\mathbb{Z}\}$ on a qubit. For $\alpha \in (0,1]$ and for integers $\alpha \geq 2$, we have

$$H_\alpha^T(X) + H_\alpha^T(Y) + H_\alpha^T(Z) \geq 2\log e \cdot f_\alpha(2)\,, \tag{364}$$

$$\text{where} \quad f_\alpha(x) := \left(\frac{1-x^{1-\alpha}}{\alpha-1}\right)\,. \tag{365}$$

This generalizes the result in (87), which is recovered by taking the limit $\alpha \to 1$, noting that $\lim_{\alpha\to 1} f_\alpha(x) = \log x/\log e$.

A more general scenario was considered in (Rastegin, 2013), where system A has dimension d, and the measurements under consideration form a set of n MUBs, $\{\mathbb{X}_j\}$. For $\alpha \in (0,2]$, Rastegin (2013) shows that

$$\frac{1}{n}\sum_{j=1}^n H_\alpha^T(X_j) \geq 2\log e \cdot f_\alpha\left(\frac{nd}{n+d-1}\right)\,. \tag{366}$$

This result is quite general in that it holds for any n and d. Furthermore, in the case of $n = d+1$ and $\alpha \to 1$, one recovers the result presented in (80). So (366) generalizes (80). Rastegin (2013) also tightened this bound for mixed states, with a state-dependent bound

$$\frac{1}{n}\sum_{j=1}^n H_\alpha^T(X_j) \geq 2\log e \cdot f_\alpha\left(\frac{nd}{n+d\,\mathrm{tr}(\rho^2)-1}\right)\,. \tag{367}$$

Other entropy families are also discussed in the literature. For example, Zozor *et al.* (2014) consider a broad class of entropies defined as

$$H_{(h,\phi)}(X) := h\left(\sum_x \phi(P_X(x))\right)\,. \tag{368}$$

Here $h : \mathbb{R} \to \mathbb{R}$ and $\phi : [0;1] \to \mathbb{R}$ are generic continuous functions such that either ϕ is strictly concave and h is strictly increasing, or ϕ is strictly convex and h is strictly decreasing. Additionally, they imposed $\phi(0) = 0$ and $h(\phi(1)) = 0$. This family includes as special cases both the Rényi and Tsallis families and hence also the Shannon entropy.

In addition to giving a detailed overview of the literature on entropic uncertainty relations, Zozor *et al.* (2014) derived a new uncertainty relation for the $H_{(h,\phi)}$ entropies. For any two POVMs $\mathbb{X}$ and $\mathbb{Z}$, and for any two pairs of functionals (h_1,ϕ_1) and (h_2,ϕ_2), their relation takes the form,

$$H_{(h_1,\phi_1)}(X) + H_{(h_2,\phi_2)}(Z) \geq B_{(h_1,\phi_1),(h_2,\phi_2)}(t)\,, \tag{369}$$

where the right-hand side is a function of the triplet

$$t := \{c_{\mathbb{X}}, c_{\mathbb{Z}}, c\}, \quad c_{\mathbb{X}} := \max_x \|\mathbb{X}^x\|, \quad c_{\mathbb{Z}} := \max_z \|\mathbb{Z}^z\|\,, \tag{370}$$

and c is defined in (49). The reader is referred to (Zozor *et al.*, 2014) for the explicit form of the bound $B_{(h_1,\phi_1),(h_2,\phi_2)}(t)$. In general, this bound can be computed, since it only involves a one-parameter optimization over a bounded interval. Note that the functionals associated with the two terms in (369) may be different. This gives a very general result allowing the authors to consider, e.g., Rényi entropy uncertainty relations that go beyond the usual conjugacy curve, defined by $(1/\alpha) + (1/\beta) = 2$.

B. Certainty relations

Instead of lower bounding sums of entropies for different observables, it is interesting to ask whether there exist non-trivial *upper* bounds on such sums. Upper bounds of this nature are called certainty relations. Of course, one would not expect to find non-trivial upper bounds for, say, the maximally mixed state $\rho_A = \mathbb{1}/d$. But suppose one restricts to pure states $|\psi\rangle_A$.

For some sets of observables, even restricting to pure states is not enough to get a certainty relation. For example, consider the standard σ_X and σ_Z observables on a qubit. One cannot find a certainty relation for these two observables because there exist states, namely the eigenstates of σ_Y, that are unbiased with respect to the eigenbases of σ_X and σ_Z, and hence lead to maximum uncertainty in these two bases. So the best bound we can get is $H(X) + H(Z) \leq 2$, which is obviously trivial.

Recently Korzekwa *et al.* (2014a) proved a general result that non-trivial certainty relations are not possible for two arbitrary orthonormal bases $\mathbb{X}$ and $\mathbb{Z}$, in any finite dimension d. This follows from the fact that one can always find a pure state $|\psi\rangle$ that is unbiased with respect to both $\mathbb{X}$ and $\mathbb{Z}$.

However, there do exist non-trivial certainty relation, e.g., for a $d+1$ set of MUBs, which is obviously connected to the fact that there are no states that are unbiased to all bases in $d+1$ MUB set. Consider a result from (Sánchez, 1993), which deals with three MUBs $(\mathbb{X}, \mathbb{Y}, \mathbb{Z})$ on a qubit system in a pure state, and says that

$$H(X) + H(Y) + H(Z) \leq \frac{3}{2}\log 6 - \frac{\sqrt{3}}{2}\log(2+\sqrt{3})\,. \tag{371}$$

The right-hand side of (371) is ≈ 2.23. Comparing this to the lower bound of 2, from (87), one sees that the allowable range for $H(X)+H(Y)+H(Z)$ is quite small. Sánchez (1993) noted that (371) is in fact the optimal certainty relation for these observables. More generally, considering a $d+1$ set of MUBs $\{\mathbb{X}_j\}$, Sánchez (1993) showed that

$$\sum_{j=1}^{n} H(X_j) \leq n\log(n+\sqrt{n}) - \frac{1}{d}(n+(n-2)\sqrt{n})\log(2+\sqrt{n})\,. \tag{372}$$

where $n = d+1$. Note that (371) is a special case of (372) corresponding to $d = 2$.

Rastegin obtained some generalizations of (371) to the Rényi and Tsallis entropy families. In the Rényi case Rastegin (2014) found, for all $\alpha \in (0,1]$,

$$H_\alpha(X) + H_\alpha(Y) + H_\alpha(Z) \leq 3R_\alpha\,, \tag{373}$$

where

$$R_\alpha := \frac{1}{1-\alpha}\log\left(\left(\frac{1+1/\sqrt{3}}{2}\right)^\alpha + \left(\frac{1-1/\sqrt{3}}{2}\right)^\alpha\right)\,. \tag{374}$$

Likewise Rastegin (2012) found a similar sort of bound for the Tsallis entropies, but with $\log(x)$ in (374) replaced by $x-1$.

While the above certainty relations are for MUBs, very recently Puchała *et al.* (2015) studied a more general situation. They considered sets of $n > 2$ orthonormal bases in dimension d and derived certainty relations for this very general situation. Their certainty relations are upper bounds on the sum of Shannon entropies, similar to (372), but are not restricted to MUBs. Certainty relations for k-designs with $k = 2, 4$ in terms of the mutual information can also be found in (Matthews *et al.*, 2009).

Finally, it is worth reminding the reader that for the collision entropy one can obtain an *equality*, as in (81). An equation of this sort is both an uncertainty and a certainty relation. Stated another way, an equation implies that the strongest uncertainty relation coincides with the strongest certainty relation, leaving no gap between the two bounds. Equations such as (81) can, in turn, be used to derive certainty relations for other entropies, such as the min-entropy, due to the fact that $H_{\min} \leq H_2$.

The generalization of (81) to bipartite states ρ_{AB} was given in (188). Equation (188) is a certainty relation in the presence of quantum memory. It relates the amount of uncertainty to the amount of entanglement, as quantified by the conditional entropy $H_2(A|B)$. Similar to the unipartite case, (188) can be used to derive certainty relations (in the presence of quantum memory) for other entropies, such as the min-entropy, as discussed in (Berta *et al.*, 2014a).

Bipartite certainty relations (i.e., in the presence of quantum memory) are an interesting open problem. For example, one could ask whether (371) or (372) can be appropriately generalized to the quantum memory case.

C. Measurement uncertainty

This review has focused on preparation uncertainty relations. Two other aspects of the uncertainty principle are (1) the joint measurability of observable pairs and (2) the disturbance of one observable caused by the measurement of another observable. Joint measurability and measurement-disturbance are two aspects of measurement uncertainty, which deals with fundamental restrictions on ones ability to measure things. For a detailed discussion of measurement uncertainty, we refer the reader to (Busch *et al.*, 2007, 2014a; Hall, 2004; Ozawa, 2003). It is important, though, that we mention measurement uncertainty here because the topic has seen significant debate recently (see, e.g., (Busch *et al.*, 2013, 2014a,b)). It represents an area where further research is needed, and entropic approaches could play a valuable role.

Rather than delve into the conceptual issues of measurement uncertainty, we will simply give a taste here of a few recent works that have taken an entropic approach, in particular, to measurement-disturbance.

1. State-independent measurement-disturbance relations

A promising approach to measurement uncertainty asks the question: how well can a measurement device perform on particular idealized sets of input states, e.g., the basis states associated with two complementary observables $\mathbb{X}$ and $\mathbb{Z}$. This is often called a state-independent approach, although it could also be called a calibration approach, since one is calibrating a device's performance based on idealized input states. For example, this approach was taken for the position and momentum observables in (Busch *et al.*, 2013), although the quantities in their relation were not entropic so we will not discuss it.

More recently the calibration approach was taken

by Buscemi *et al.* (2014). They considered the sequential measurement (or more generally, a joint measurement) of two orthonormal bases, $\mathbb{X}$ and $\mathbb{Z}$, and employed entropic measures of $\mathbb{X}$ measurement noise and $\mathbb{Z}$ disturbance. Bob's measurement of Alice's system A was modeled using a quantum channel $\mathcal{M}$ that maps A to two classical registers M_X and M_Z that respectively store Bob's outcomes for his attempted $\mathbb{X}$ and $\mathbb{Z}$ measurements. $\mathcal{M}$ can be decomposed into complementary quantum channels $\mathcal{E}$ and $\mathcal{F}$ (see (255) for the definition of complementary channels), where $\mathcal{E}$ maps A to the classical register M_X. The complementary channel $\mathcal{F}$ maps A to a bipartite system $A'M'_X$, where A' is simply system A at a later time, and M'_X is an additional copy of the classical register M_X. (When we say a register is "classical", we mean that there exists another copy of it somewhere, and in this case M'_X is that copy.) Finally, Bob uses a recovery map $\mathcal{R}$ to map $A'M'_X$ to the classical register M_Z.

In the spirit of the calibration approach, they supposed that Alice sends the $\mathbb{X}$ basis states with equal probability into the channel $\mathcal{E}$, and quantified the $\mathbb{X}$ measurement noise by $\mathsf{N}(\mathcal{M},\mathbb{X}) = H(X_A|M_X)$. Alternatively, they supposed Alice sends the $\mathbb{Z}$ basis states with equal probability into the channel $\mathcal{R}\circ\mathcal{F}$, and quantified the disturbance of $\mathbb{Z}$ by $\mathsf{D}(\mathcal{M},\mathbb{Z}) = H(Z_A|M_Z)$. Here X_A and Z_A refer to Alice's classical registers that store which basis state she sent. Their measurement-disturbance relation states that

$$\mathsf{N}(\mathcal{M},\mathbb{X}) + \mathsf{D}(\mathcal{M},\mathbb{Z}) \geq \log\frac{1}{c}\,, \tag{375}$$

which shows a trade-off between Bob's ability to measure the $\mathbb{X}$ states well versus his ability to leave the $\mathbb{Z}$ states undisturbed. Buscemi *et al.* (2014) emphasize that their measures have a close connection to the guessing probability, and hence are well-motivated operationally. Furthermore, their allowance of a recovery map $\mathcal{R}$ captures the notion of irreversible disturbance.

Formally speaking, we remark that (375) can be derived from (257), which deals with sending information over complementary quantum channels. One can see this using

$$\mathsf{N}(\mathcal{M},\mathbb{X}) = \log d - \chi(\mathcal{E},\mathbb{X}) \tag{376}$$

$$\mathsf{D}(\mathcal{M},\mathbb{Z}) = \log d - \chi(\mathcal{R}\circ\mathcal{F},\mathbb{Z})\,. \tag{377}$$

These identities allow one to convert between the forms in (375) and (257).

Recently Renes and Scholz (2014) introduced an alternative approach that is not explicitly entropic, but it has both a clear operational interpretation and relevance to information-processing tasks. Namely they quantify error and disturbance in terms of the probability of distinguishing the channel associated with the actual measurement apparatus from a channel that performs an ideal measurement.

2. State-dependent measurement-disturbance relations

Now let us consider a sequential measurement scenario where Alice feeds in system A in an arbitrary state ρ_A. We now discuss some state-dependent relations, i.e., relations that hold for any input state ρ_A.

For simplicity, consider the sequential measurement of orthonormal bases, $\mathbb{X}$ followed by $\mathbb{Z}$, where the first measurement is a von Neumann measurement, i.e., it leaves the system in an $\mathbb{X}$ basis state. One can apply Maassen-Uffink's uncertainty relation to each outcome of the $\mathbb{X}$ measurement, i.e., to each state $|\mathbb{X}^x\rangle$, giving

$$H(Z)_{|\mathbb{X}^x\rangle} = H(X)_{|\mathbb{X}^x\rangle} + H(Z)_{|\mathbb{X}^x\rangle} \geq \log\frac{1}{c}\,. \tag{378}$$

Multiplying this by the probability $p^x = \langle\mathbb{X}^x|\rho_A|\mathbb{X}^x\rangle$ for outcome x, and summing over x gives

$$H(Z|X) \geq \log\frac{1}{c}\,, \tag{379}$$

where $H(Z|X)$ denotes the uncertainty for a future $\mathbb{Z}$ measurement given the outcome of the previous $\mathbb{X}$ measurement. Equation (379) was discussed in detail by Baek *et al.* (2014), and was also briefly mentioned by Coles and Piani (2014a). Note that (379) holds for any input state ρ_A, so technically it is a state-dependent relation.

While (379) assumes the $\mathbb{X}$ measurement is an ideal von Neumann measurement, it is interesting to ask what happens if the first measurement is non-ideal. There are various ways to address this. One approach, given by Coles and Furrer (2015), quantified the imperfection of the $\mathbb{X}$ measurement by the predictive error,

$$\mathsf{E}(\rho_A,\mathbb{X},\mathcal{E}) := H_{\max}(X|M_X)\,, \tag{380}$$

that is, the max-entropy of a future (perfect) $\mathbb{X}$ measurement given the register M_X that stores the outcome of the previous (imperfect) measurement of $\mathbb{X}$. Here $\mathcal{E}$, which maps $A \to AM_X$, is the channel that performs this imperfect $\mathbb{X}$ measurement. One is interested in the disturbance of the $\mathbb{Z}$ observables caused by the imperfect $\mathbb{X}$ measurement. Coles and Furrer (2015) quantified the disturbance of $\mathbb{Z}$ using the Rényi relative entropies for $\alpha \in [1/2,\infty]$,

$$\mathsf{D}_\alpha(\rho_A,\mathbb{Z},\mathcal{E}) := D_\alpha(P_Z||P_Z^{\mathcal{E}})\,. \tag{381}$$

Here P_Z is the initial probability distribution for the $\mathbb{Z}$ measurement and $P_Z^{\mathcal{E}}$ is the final probability distribution for $\mathbb{Z}$, i.e., after the imperfect $\mathbb{X}$ measurement. With these definitions, they found the measurement-disturbance relation

$$\mathsf{D}_\alpha(\rho_A,\mathbb{Z},\mathcal{E}) + \mathsf{E}(\rho_A,\mathbb{X},\mathcal{E}) + H_\alpha(Z)_P \geq \frac{1}{c}\,. \tag{382}$$

On the one hand this gives a trade-off between measuring $\mathbb{X}$ well and causing large $\mathbb{Z}$ disturbance. On the

other hand, the trade-off gets weaker as more initial uncertainty is contained in P_Z, as quantified by the term $H_\alpha(Z)_P$. So there is an interplay between initial uncertainty, measurement error, and disturbance.

VIII. PERSPECTIVES

We have discussed modern formulations of Heisenberg's uncertainty principle, where uncertainty is quantified by entropy. Such formulations are directly relevant to quantum information processing tasks as discussed in Sec. VI.

Technological applications such as QKD (Sec. VI.B) provide the driving force for obtaining better entropic uncertainty relations. For example, to prove security of QKD protocols involving more than two measurements, new entropic uncertainty relations are needed - namely ones that allow for quantum memory and for multiple measurements. This is an important frontier that requires more research. Device-independent randomness, i.e., certifying randomness obtained from untrusted devices (Sec. VI.A.2), is another emerging application for which entropic uncertainty relations appear to be useful but more research is needed to find uncertainty relations that are specifically tailored to this application.

Aside from their technological applications, we believe entropic uncertainty relations have a beauty to them. They give insight into the structure of quantum theory, and for that reason alone they are worth pursuing. For example Sec. IV.F.5 noted a simple conjecture — that the sum of the mutual informations for two MUBs lower bounds the quantum mutual information. It would be interesting to prove this conjecture because of its beautiful simplicity.

New tools are now being developed to prove entropic uncertainty relations. The majorization approach (Sec. III.I) is quite promising and very recently it has been extended to allow for a memory system (Gour *et al.*, 2015). The relation between the majorization proof approach and the relative-entropy proof approach (e.g., see App. B) remains to be clarified, and a unified framework would be insightful. For uncertainty relations with memory, Dupuis *et al.* (2015) establish a powerful meta theorem to derive new uncertainty relations. Yet, it is known that the resulting relations are not tight in all regimes, calling for further improvements.

One of most exciting things about entropic uncertainty relations is that they are showing promise for giving insight into basic physics. For example Sec VI.F discussed how entropic uncertainty relations allow one to unify the uncertainty principle with the wave-particle duality principle. A natural framework for quantifying wave-particle duality will likely come from applying entropic uncertainty relations to interferometers. Likewise, a hot topic in quantum foundations is measurement uncertainty. Sec. VII.C noted that entropic uncertainty relations may play an important role in obtaining conceptually clear formulations of measurement uncertainty. Furthermore, entropic uncertainty relations will continue to help researchers characterize the boundary between separable vs. entangled states (Sec. VI.D), as well as steerable vs. non-steerable states (Sec. VI.E).

But other physics applications are emerging. Entropic uncertainty relations may play a role in the study of phase transitions in condensed matter physics (Romera and Calixto, 2015). Entropic uncertainty relations are also studied in the context of special and general relativity (Feng *et al.*, 2013; Jia *et al.*, 2015). Given that quantum information is playing an increasing role in cosmology (Hayden and Preskill, 2007), it would not be surprising to see future work on entropic uncertainty relations in the context of black hole physics.

ACKNOWLEDGMENTS

We thank Kais Abdelkhalek, Iwo Białynicki-Birula, Matthias Christandl, Rupert L. Frank, Gilad Gour, Michael J. W. Hall, Hans Maassen, Joseph M. Renes, Renato Renner, Lukasz Rudnicki, Christian Schaffner, Reinhard F. Werner, Mark M. Wilde, and Karol Zyczkowski for feedback. PJC acknowledges support from Industry Canada, Sandia National Laboratories, NSERC Discovery Grant, and Ontario Research Fund (ORF). MB acknowledges funding provided by the Institute for Quantum Information and Matter, an NSF Physics Frontiers Center (NFS Grant PHY-1125565) with support of the Gordon and Betty Moore Foundation (GBMF-12500028). Additional funding support was provided by the ARO grant for Research on Quantum Algorithms at the IQIM (W911NF-12-1-0521). MT is funded by an University of Sydney Postdoctoral Fellowship and acknowledges support from the ARC Centre of Excellence for Engineered Quantum Systems (EQUS). SW is supported by STW, Netherlands and an NWO VIDI grant.

Appendix A: Mutually unbiased bases

Sec. III.B.2 defined MUBs, and sets of n MUBs. The study of MUBs is closely related to the study of entropic uncertainty. Strong entropic uncertainty relations have been derived generically for sets of MUBs (particularly for $d+1$ sets of MUBs), hence constructing a new set of MUBs immediately yields a new entropic uncertainty relation. On the other hand, there are interesting open questions about whether a set of n MUBs $\{\mathbb{X}_j\}$ yield the strongest bound b in a generic uncertainty relation of the

form

$$\sum_{j=1}^{n} H(X_j) \geq b\,. \tag{A1}$$

A review of MUBs can be found in (Durt *et al.*, 2010). Here we discuss the connection of MUBs to Hadamard matrices, as well as the existence and construction of MUBs.

1. Connection to Hadamard matrices

Any two orthonormal bases are related by a unitary, and in the case of MUBs, that unitary is called a Hadamard matrix H. The general form of such matrices is

$$H = \sum_{j,k} \frac{e^{i\phi_{jk}}}{\sqrt{d}} |j\rangle\langle k|\,, \tag{A2}$$

where the phase factors ϕ_{jk} must be appropriately chosen so that H is unitary. Notice that each matrix element has a magnitude of $1/\sqrt{d}$, which is the defining property of Hadamard unitaries. The most well-known Hadamard is the Fourier matrix, defined in (207) but restated here,

$$F = \sum_{j,k} \frac{\omega^{-jk}}{\sqrt{d}} |j\rangle\langle k| \quad \text{with} \quad \omega = e^{2\pi i/d}\,, \tag{A3}$$

which relates the generalized Pauli operators

$$\sigma_{\mathbb{Z}} = \sum_j \omega^j |j\rangle\langle j|, \quad \sigma_{\mathbb{X}} = F\sigma_{\mathbb{Z}}F^\dagger = \sum_j |j+1\rangle\langle j|\,. \tag{A4}$$

For $d = 2$ these are just the usual Pauli matrices from Ex. 7.

It should be clear that the problem of finding MUBs is equivalent to the problem of finding Hadamard matrices. We note that Hadamards can be categorized into equivalence classes, based on whether there exists a diagonal unitary or permutation that that maps one Hadamard to another. A detailed catalog of Hadamard matrices can be found online (Bruzda *et al.*, 2015).

2. Existence

That there exist MUB pairs in any finite dimension follows, e.g., from the fact that we can write down F in (A3) for any d. In fact, for any d there exists a set of 3 MUBs, e.g., formed from the eigenvectors of $\sigma_{\mathbb{X}}$, $\sigma_{\mathbb{Z}}$, and $\sigma_{\mathbb{X}}\sigma_{\mathbb{Z}}$.

It is also known that a set of MUBs can at most be of size $d+1$ (Bandyopadhyay *et al.*, 2002). Such $d+1$ sets (if they exist) are called complete sets of MUBs. Complete sets play a role in tomography since they are informationally complete, and they have the useful property of forming a complex projective two-design (Klappenecker and Rotteler, 2005).

Complete sets of MUBs are known to exist in prime power dimensions, i.e., $d = p^m$ where p is a prime and m is a positive integer. However, even for the smallest number that is not a prime power, namely 6, the existence problem remains unsolved.

3. Simple constructions

When d is a prime, a simple construction (Bandyopadhyay *et al.*, 2002; Wootters and Fields, 1989) of a complete set of MUBs is to consider the eigenvectors of the $d+1$ products of the form

$$\{\sigma_{\mathbb{Z}}, \sigma_{\mathbb{X}}, \sigma_{\mathbb{X}}\sigma_{\mathbb{Z}}, \sigma_{\mathbb{X}}\sigma_{\mathbb{Z}}^2, ..., \sigma_{\mathbb{X}}\sigma_{\mathbb{Z}}^{d-1}\}. \tag{A5}$$

More generally for $d = p^m$, a construction is known where each basis B_i comes from the common eigenvectors of a corresponding set C_i of commuting matrices (Bandyopadhyay *et al.*, 2002). The elements of C_i are a subset of size $|C_i| = d-1$ of the d^2-1 Pauli products $\sigma_{\mathbb{X}}^j\sigma_{\mathbb{Z}}^k$ (excluding the identity). The subset is chosen such that all the elements of C_i commute and $C_i \bigcap C_j = \{\mathbb{1}\}$ for $i \neq j$.

Appendix B: Proof of Maassen-Uffink's Shannon entropy relation

Here we give a simple proof of Maassen-Uffink's uncertainty relation for the Shannon entropy (31). Our proof closely follows the ideas in (Coles *et al.*, 2012) and makes use of the data-processing inequality for Umegaki's relative entropy (Lindblad, 1975; Uhlmann, 1977). In fact, we will prove the slightly stronger relation stated in (47) and repeated here,

$$H(X) + H(Z) \geq \log\frac{1}{c} + H(A)\,. \tag{B1}$$

Note that this strengthens the Maassen-Uffink relation (31) when the state ρ_A has some mixedness.

Proof. For the proof of (B1) we consider the classical state $\tilde{\rho}_X = \mathcal{X}_{A\to X}(\rho_A)$ generated by applying the measurement map

$$\mathcal{X}_{A\to X}(\cdot) = \sum_x \langle \mathbb{X}^x| \cdot |\mathbb{X}^x\rangle |x\rangle\langle x|_X\,, \tag{B2}$$

where $\{|x\rangle\}_x$ is an orthonormal basis of an auxiliary Hilbert space X that allows us to represent the classical random variable X in the quantum formalism.

It is easy to verify that the Shannon entropy of the distribution P_X is equal to the von Neumann entropy of the state $\tilde{\rho}_X$. From this we get

$$H(X) = -\operatorname{tr}[\tilde{\rho}_X \log \tilde{\rho}_X] = -\operatorname{tr}[\mathcal{X}(\rho_A)\log\mathcal{X}(\rho_A)] \tag{B3}$$

$$= -\operatorname{tr}[\rho_A \log \mathcal{X}(\rho_A)]\,, \tag{B4}$$

where the last equality is straightforward to check by writing out the trace and the measurement map $\mathcal{X}_{A\to X}$. By phrasing the right-hand side of (B3) in terms of Umegaki's relative entropy,

$$D(\rho_A\|\sigma_A) = \mathrm{tr}[\rho_A(\log\rho_A - \log\sigma_A)]\,, \tag{B5}$$

we arrive at

$$H(X) = D(\rho_A\|\mathcal{X}(\rho_A)) + H(A)_\rho\,. \tag{B6}$$

We then apply the measurement map

$$\mathcal{Z}_{A\to Z}(\cdot) = \sum_z \langle\mathbb{Z}^z|\cdot|\mathbb{Z}^z\rangle|z\rangle\langle z|_Z\,. \tag{B7}$$

to both arguments of the relative entropy, and find by the data-processing inequality for the relative entropy (Lindblad, 1975; Uhlmann, 1977) that

$$\begin{aligned} D(\rho_A\|\mathcal{X}(\rho_A)) &\geq D(\mathcal{Z}(\rho_A)\|\mathcal{Z}\circ\mathcal{X}(\rho_A)) \\ &= D(\tilde\rho_Z\|\mathcal{Z}\circ\mathcal{X}(\rho_A))\,, \end{aligned} \tag{B8}$$

where $\tilde\rho_Z = \mathcal{Z}_{A\to Z}(\rho_A)$. By writing out both measurement maps we find the classical state

$$\mathcal{Z}\circ\mathcal{X}(\rho_A) = \sum_z |z\rangle\langle z|_Z \cdot \sum_x |\langle\mathbb{X}^x|\mathbb{Z}^z\rangle|^2\langle\mathbb{X}^x|\rho_A|\mathbb{X}^x\rangle\,, \tag{B9}$$

and the relative entropy on the right-hand side of (B8) becomes

$$\begin{aligned} D(\tilde\rho_Z\|\mathcal{Z}\circ\mathcal{X}(\rho_A)) &= -H(Z)_{\tilde\rho} \\ &\quad - \sum_z \langle\mathbb{Z}^z|\rho_A|\mathbb{Z}^z\rangle \log\left(\sum_x |\langle\mathbb{X}^x|\mathbb{Z}^z\rangle|^2\langle\mathbb{X}^x|\rho_A|\mathbb{X}^x\rangle\right). \end{aligned} \tag{B10}$$

Now, the logarithm is a monotonic function and hence we can bound

$$\begin{aligned} &- \sum_z \langle\mathbb{Z}^z|\rho_A|\mathbb{Z}^z\rangle \log\left(\sum_x |\langle\mathbb{X}^x|\mathbb{Z}^z\rangle|^2\langle\mathbb{X}^x|\rho_A|\mathbb{X}^x\rangle\right) \\ &\geq - \sum_z \langle\mathbb{Z}^z|\rho_A|\mathbb{Z}^z\rangle \log\left(\max_{x',z'} |\langle\mathbb{X}^{x'}|\mathbb{Z}^{z'}\rangle|^2 \sum_x \langle\mathbb{X}^x|\rho_A|\mathbb{X}^x\rangle\right) \end{aligned} \tag{B11}$$

$$= -\log \max_{x',z'} |\langle\mathbb{X}^{x'}|\mathbb{Z}^{z'}\rangle|^2\,. \tag{B12}$$

By combining (B3)–(B12) and noting that $H(Z)$ equals the von Neumann entropy of $\tilde\rho_Z$, we arrive at the claim (B1). □

Appendix C: Rényi entropies for joint quantum systems

We will now define conditional Rényi entropies for general quantum states. This allows us to exhibit their intuitive properties in a very general setting without having to discuss various special cases individually. We will exhibit these properties to show a generalization of the Maassen-Uffink relation to the tripartite quantum memory setting.

1. Definitions

For any bipartite quantum state ρ_{AB} and $\alpha \in [\frac{1}{2}, \infty]$, we define the *quantum conditional Rényi entropy* as

$$H_\alpha(A|B) := -\min_{\sigma_B} D_\alpha(\rho_{AB}\|\mathbb{1}_A\otimes\sigma_B)\,, \quad \text{where } \sigma_B \text{ is a quantum state on } B\,. \tag{C1}$$

Here, D_α is the Rényi divergence of order α (Müller-Lennert *et al.*, 2013; Wilde *et al.*, 2014), namely[30]

$$D_\alpha(\rho\|\sigma) := \frac{1}{\alpha-1}\log\mathrm{tr}\left[\left(\sigma^{\frac{1-\alpha}{2\alpha}}\rho\sigma^{\frac{1-\alpha}{2\alpha}}\right)^\alpha\right] \quad \text{for } \alpha\in\left[\frac{1}{2},1\right)\cup(1,\infty) \tag{C2}$$

and as the corresponding limit for $\alpha \in \{1, \infty\}$. These divergences are measures of distinguishability between quantum states and some of their properties will be discussed in App. C.2.

Here we want to note the following special cases that we have encountered previously. First, the conditional min- and max-entropies are simply recovered as $H_{\min} \equiv H_\infty$ and $H_{\max} \equiv H_{1/2}$. The von Neumann entropy is recovered as $H \equiv H_1$. Finally, the conditional collision entropy can be expressed as

$$H_{\mathrm{coll}}(A|B) = -D_2(\rho_{AB}\|\mathbb{1}_A\otimes\rho_B)\,. \tag{C3}$$

Note that $H_2(A|B) \leq H_{\mathrm{coll}}(A|B)$ since the former involves a minimization over marginal states σ_B. The two expressions are not equal in general and we want to work with $H_{\mathrm{coll}}(A|B)$ because it has the operational interpretation as in (141) and (150).

2. Entropic properties

We present the properties for the whole family of Rényi divergences and entropies here, but recall that the properties also apply to the relative entropy and the von Neumann entropy as special cases. Most properties of the conditional Rényi entropy can be derived from properties of the underlying Rényi divergence.[31]

---

[30] This quantum generalization is not unique — in fact other generalizations based on Petz's notion of Rényi divergence (Ohya and Petz, 1993) have also been explored, for example in (Tomamichel *et al.*, 2014). However, for the purpose of the present review it is convenient to stick with the proposed general definition in (C1) and (C2) as it entails the most important special cases encountered here and in the literature.

[31] These divergences have been investigated in a series of recent works (Beigi, 2013; Frank and Lieb, 2013b; Mosonyi and Ogawa, 2015; Müller-Lennert *et al.*, 2013; Wilde *et al.*, 2014) and proofs of the properties discussed here can be found in these references.

a. Positivity and monotonicity

First, we remark that $D_\alpha(\rho\|\sigma)$ is guaranteed to be non-negative when the arguments ρ and σ are normalized, and $D_\alpha(\rho\|\sigma) = 0$ when $\rho = \sigma$. Also, $\alpha \mapsto D_\alpha(\rho\|\sigma)$ is monotonically increasing in α. Thus, for any $\beta \geq \alpha$, we have

$$\begin{aligned} 0 \leq D_\alpha(\rho\|\sigma) \leq D_\beta(\rho\|\sigma)\,, \quad \text{and} \\ \log d_A \geq H_\alpha(A|B)_\rho \geq H_\beta(A|B)_\rho \geq -\log\min\{d_A, d_B\}\,. \end{aligned} \tag{C4}$$

This means that the conditional Rényi entropies, in particular also the von Neumann entropy, can be negative. However, this can only happen in the presence of quantum entanglement and the conditional entropies are thus always positive when one of the two systems is classical.

The maximum, $\log d_A$, is achieved for a state of the form $\rho_{AB} = \tau_A \otimes \rho_B$ where τ_A is fully mixed. On the other hand, the minimum, $-\log d_A$, is achieved for a maximally entangled pure state of the form $|\psi\rangle_{AB} = \frac{1}{\sqrt{d_A}}\sum_x |x\rangle_A \otimes |x\rangle_B$.

b. Data-processing inequalities

The Rényi divergences satisfy a data-processing inequality. Namely, for all $\alpha \geq \frac{1}{2}$ and any cptp map $\mathcal{E}$, we find the following relation (Frank and Lieb, 2013b):

$$D_\alpha\big(\mathcal{E}(\rho)\big\|\mathcal{E}(\sigma)\big) \leq D_\alpha(\rho\|\sigma)\,. \tag{C5}$$

This is an expression of the intuitive property that it is easier to distinguish between the inputs rather than the outputs of any quantum channel. In fact, this property holds more generally for any completely positive trace non-increasing map $\mathcal{E}$ which satisfies $\mathrm{tr}[\mathcal{E}(\rho)] = 1$. This property has two important implications for conditional entropies.

First, consider an arbitrary cptp map $\mathcal{E}_{B\to B'}$ acting on the side information that takes ρ_{AB} to $\tau_{AB'} = \mathcal{I}_A \otimes \mathcal{E}_{B\to B'}(\rho_{AB})$. Then we have $H_\alpha(A|B)_\rho \leq H_\alpha(A|B')_\tau$. This tells us that any physically allowed information processing of the side information B may only increase the uncertainty we have about A.

Example 34. *An often encountered special case of this is the inequality $H_\alpha(A|BC)_\rho \leq H_\alpha(A|B)_\rho$ for any tripartite state ρ_{ABC}, which expresses the fact that throwing away part of the side information can only increase the uncertainty about A.*

The second application concerns rank-1 projective measurements on the A system. More precisely, we consider any rank-1 projective measurement $\mathcal{X}_{A\to X}$ that takes ρ_{AB} to

$$\begin{aligned} \tilde{\rho}_{XB} &= \mathcal{X}_{A\to X} \otimes \mathcal{I}_B(\rho_{AB}) \\ &= \sum_x |x\rangle\langle x|_X \cdot \Big(\langle \mathbb{X}^x|_A \otimes \mathbb{1}_B\Big)\rho_{AB}\Big(|\mathbb{X}^x\rangle_A \otimes \mathbb{1}_B\Big)\,. \end{aligned} \tag{C6}$$

Then, we find that $H_\alpha(A|B)_\rho \leq H_\alpha(X|B)_\tau$, which reveals that measuring out system A completely can only increase the uncertainty we have about it.[32]

c. Duality and additivity

We will see that the following property is essential for deriving uncertainty relations with quantum side information. For any tripartite state ρ_{ABC}, the conditional Rényi entropies satisfy the following *duality relation*. For $\alpha, \beta \in [\frac{1}{2}, \infty]$ such that $\frac{1}{\alpha} + \frac{1}{\beta} = 2$, we have (Beigi, 2013; Müller-Lennert *et al.*, 2013)

$$H_\alpha(A|B)_\rho + H_\beta(A|C)_\rho \geq 0\,, \tag{C7}$$

with equality if ρ_{ABC} is pure.

This is a quantitative manifestation of the monogamy of quantum correlations. For example, if system A is highly entangled with system B we find that the conditional von Neumann entropy $H(A|B)_\rho$ is negative. However, the duality relation (C7) now shows that for any third system C correlated with A and B, it holds that $H(A|C)_\rho \geq -H(A|B)_\rho$, that is, the uncertainty of A from an observer with access to C is necessarily large in this case.

The Rényi entropies are additive. Namely, given a product state of the form $\rho_{ABCD} = \rho_{AC} \otimes \rho_{BD}$, they satisfy $H_\alpha(AB|CD)_\rho = H_\alpha(A|C)_\rho + H_\alpha(B|D)_\rho$. This is in fact a consequence of the above duality relation.[33]

3. Axiomatic proof of uncertainty relation with quantum memory

Here we give a concise proof of the generalized Maassen-Uffink relation (225), restated here

$$H_\alpha(X|B)_\rho + H_\beta(Z|C)_\rho \geq \log\frac{1}{c}\,, \tag{C11}$$

---

[32] The above inequality holds a little more generally for all cptp maps on $\mathcal{E}_{A\to A'}$ that satisfy $\mathcal{E}_{A\to A'}(\mathbb{1}_A) = \mathbb{1}_{A'}$, e.g., unital maps.

[33] Recall that by definition (C1), we have

$$\begin{aligned} H_\alpha(AB|CD)_\rho &= -\min_{\sigma_{CD}} D_\alpha(\rho_{ABCD}\|\mathbb{1}_{AB}\otimes\sigma_{CD}) && \text{(C8)} \\ &\geq -\min_{\sigma_C,\sigma_D} D_\alpha(\rho_{ABCD}\|\mathbb{1}_{AB}\otimes\sigma_C\otimes\sigma_D) && \text{(C9)} \\ &= H_\alpha(A|C)_\rho + H_\alpha(B|D)_\rho\,. && \text{(C10)} \end{aligned}$$

The reverse inequality then follows due to the duality relation.

where $\frac{1}{\alpha}+\frac{1}{\beta}=2$. Let us note that the proof applies to a very general class of entropic quantities that satisfy certain properties, but we will specialize it here to conditional Rényi entropies.

Let us consider measurements $\mathbb{X}=\{\mathbb{X}_A^x\}$ and $\mathbb{Z}=\{\mathbb{Z}_A^z\}$ in two orthonormal bases such that $\mathbb{X}_A^x$ and $\mathbb{Z}_A^z$ are rank-one projectors. The more general proof for POVMs (and more general entropy measures) follows essentially the same steps and is detailed in (Coles *et al.*, 2012), based on ideas of Coles *et al.* (2011) and Tomamichel and Renner (2011).

Proof of (C11). First, let us define the isometry $V:=\sum_z |z\rangle_Z\otimes\mathbb{Z}_A^z$ associated with the $\mathbb{Z}$ measurement on system A, and the state $\tilde{\rho}_{ZABC}:=V\rho_{ABC}V^\dagger$. We find the following sequence of inequalities (which will be explained in detail below),

$$\begin{aligned}
&H_\beta(Z|C)_\rho\\
&\geq -H_\alpha(Z|AB)_\rho && \text{(C12)}\\
&=\min_{\sigma_{AB}} D_\alpha\big(\tilde{\rho}_{ZAB}\big\|\mathbb{1}_Z\otimes\sigma_{AB}\big) && \text{(C13)}\\
&\geq \min_{\sigma_{AB}} D_\alpha\left(\rho_{AB}\middle\|\sum_z \mathbb{Z}_A^z\sigma_{AB}\mathbb{Z}_A^z\right) && \text{(C14)}\\
&\geq \min_{\sigma_{AB}} D_\alpha\left(\overline{\rho}_{XB}\middle\|\sum_{x,z}\left|\langle\mathbb{X}_A^x|\mathbb{Z}_A^z\rangle\right|^2\mathbb{X}_A^x\otimes\mathrm{tr}_A\left[\mathbb{Z}_A^z\sigma_{AB}\right]\right). && \text{(C15)}
\end{aligned}$$

where we have used $\overline{\rho}_{XB}:=\sum_k \mathbb{X}_A^x\rho_{AB}\mathbb{X}_A^x$. To establish (C12), we apply the duality relation (C7) to the state $\tilde{\rho}_{ZABC}$. Equation (C13) is simply the definition of the conditional entropy as in (C1). To find (C14), we apply the data-processing inequality for the partial isometry $V^\dagger$ as a trace non-increasing map, and note that $V^\dagger(\mathbb{1}_Z\otimes\sigma_{AB})V=\sum_z\mathbb{Z}_A^z\sigma_{AB}\mathbb{Z}_A^z$. Next, (C15) follows by applying the data-processing inequality for the measurement cptp map $\mathcal{X}[\cdot]=\sum_x\mathbb{X}^x\cdot\mathbb{X}^x$.

Next we observe that

$$\begin{aligned}
&\sum_{x,z}\left|\langle\mathbb{X}_A^x|\mathbb{Z}_A^z\rangle\right|^2\mathbb{X}_A^x\otimes\mathrm{tr}_A\left[\mathbb{Z}_A^z\sigma_{AB}\right]\\
&\leq c\sum_{x,z}\mathbb{X}_A^x\otimes\mathrm{tr}_A\left[\mathbb{Z}_A^z\sigma_{AB}\right]=c\,\mathbb{1}_A\otimes\sigma_B, && \text{(C16)}
\end{aligned}$$

where we recall that $c=\max_{x,z}\left|\langle\mathbb{X}_A^x|\mathbb{Z}_A^z\rangle\right|^2$ as defined in (32). Moreover, we need that for any σ' and positive λ such that $\sigma\leq\lambda\sigma'$, we have $D_\alpha(\rho\|\sigma)\geq D_\alpha(\rho\|\sigma')+\log\frac{1}{\lambda}$.[34] Continuing from (C15), we thus find that

$$\begin{aligned}
H_\beta(Z|C)_\rho &\geq \min_{\sigma_B} D_\alpha\big(\overline{\rho}_{XB}\big\|\mathbb{1}_X\otimes\sigma_B\big)+\log\frac{1}{c} && \text{(C17)}\\
&=-H_\alpha(X|B)_\rho+\log\frac{1}{c}\,, && \text{(C18)}
\end{aligned}$$

where (C18) again follows by the definition of the conditional entropy. □

[34] For a proof of this property, see (Müller-Lennert *et al.*, 2013, Prop. 4).

REFERENCES

Abdelkhalek, K., R. Schwonnek, H. Maassen, F. Furrer, J. Duhme, P. Raynal, B.-G. Englert, and R. F. Werner (2015), "Optimality of Entropic Uncertainty Relations," arXiv:1509.00398.

Adamczak, R., R. Latała, Z. Puchała, and K. Życzkowski (2014), "Asymptotic Entropic Uncertainty Relations," arXiv:1412.7065.

Ambainis, A. (2010), Quantum Information and Computation **10** (9&10), 0848.

Azarchs, A. (2004), *Entropic Uncertainty Relations for Incomplete Sets of Mutually Unbiased Observables*, Semester thesis (California Institute of Technology).

Baek, K., T. Farrow, and W. Son (2014), Physical Review A **89** (3), 032108.

Ballester, M., and S. Wehner (2007), Physical Review A **75** (2), 022319.

Ballester, M. A., S. Wehner, and A. Winter (2008), IEEE Transactions on Information Theory **54** (9), 4183.

Bandyopadhyay,, Boykin, V. Roychowdhury, and Vatan (2002), Algorithmica **34** (4), 512.

Barnum, H., and E. Knill (2002), Journal of Mathematical Physics **43** (5), 2097.

Baumgratz, T., M. Cramer, and M. B. Plenio (2014), Physical Review Letters **113** (14), 140401.

Beckner, W. (1975), Annals of Mathematics **102** (1), 159.

Beigi, S. (2013), Journal of Mathematical Physics **54** (12), 122202.

Bennett, C. H., and G. Brassard (1984), in *Proc. IEEE International Conference on Computers, Systems and Signal Processing 1984*, Vol. 1 (IEEE, Bangalore) pp. 175–179.

Berta, M. (2013), *Quantum Side Information: Uncertainty Relations, Extractors, Channel Simulations*, Ph.D. thesis (ETH Zurich).

Berta, M., F. G. S. L. Brandao, M. Christandl, and S. Wehner (2013), IEEE Transactions on Information Theory **59** (10), 6779.

Berta, M., M. Christandl, R. Colbeck, J. M. Renes, and R. Renner (2010), Nature Physics **6** (9), 659.

Berta, M., P. J. Coles, and S. Wehner (2014a), Physical Review A **90** (6), 062127.

Berta, M., O. Fawzi, and S. Wehner (2014b), IEEE Transactions on Information Theory **60** (2), 1168.

Bialynicki-Birula, I., and L. Rudnicki (2011), in *Statistical Complexity*, edited by K. Sen (Springer Netherlands, Dordrecht) pp. 1–34.

Białynicki-Birula, I. (1984), Physics Letters A **103** (5), 253.

Białynicki-Birula, I. (2006), Physical Review A **74** (5), 10.1103/PhysRevA.74.052101.

Białynicki-Birula, I. (2007), in *AIP Conference Proceedings*, Vol. 889 (AIP, Vaxjo) pp. 52–61.

Białynicki-Birula, I., and J. L. Madajczyk (1985), Physics Letters A **108** (8), 384.

Białynicki-Birula, I., and J. Mycielski (1975), Communications in Mathematical Physics **44** (2), 129.

Biham, E., M. Boyer, P. O. Boykin, T. Mor, and V. Roychowdhury (2000), in *Proc. ACM STOC 2000* (ACM Press, New York, NY) pp. 715–724.

Biham, E., M. Boyer, P. O. Boykin, T. Mor, and V. Roychowdhury (2006), Journal of Cryptology **19** (4), 381.

Boltzmann, L. (1872), in *Sitzungsberichte der Akademie der Wissenschaften zu Wien*, Vol. 66, pp. 275–370.

Broadbent, A., and C. Schaffner (2015), "Quantum Cryptography Beyond Quantum Key Distribution," arXiv:1510.06120.

Brukner, Č., and A. Zeilinger (1999), Physical Review Letters **83** (17), 3354.

Bruß, D. (1998), Physical Review Letters **81** (14), 3018.

Bruzda, W., W. Tadej, and K. Życzkowski (2015), "Complex Hadamard Matrices," .

Buhrman, H., M. Christandl, P. Hayden, H.-K. Lo, and S. Wehner (2008), Physical Review A **78** (2), 022316.

Buscemi, F., M. J. W. Hall, M. Ozawa, and M. M. Wilde (2014), Physical Review Letters **112** (5), 050401.

Busch, P., T. Heinonen, and P. Lahti (2007), Physics Reports **452** (6), 155.

Busch, P., P. Lahti, and R. F. Werner (2013), Physical Review Letters **111** (16), 160405.

Busch, P., P. Lahti, and R. F. Werner (2014a), Review of Modern Physics **86** (4), 1261.

Busch, P., P. Lahti, and R. F. Werner (2014b), Physical Review A **89** (1), 012129.

Cachin, C., and U. Maurer (1997), in *Proc. CRYPTO 1997*, LNCS, Vol. 1294 (Springer) pp. 292–306.

Canetti, R. (2001), in *Proc. IEEE FOCS 2001* (IEEE) pp. 136–145.

Cavalcanti, E. G., S. J. Jones, H. M. Wiseman, and M. D. Reid (2009), Physical Review A **80** (3), 032112.

Cerf, N., M. Bourennane, A. Karlsson, and N. Gisin (2002), Physical Review Letters **88** (12), 127902.

Chen, B., and S.-M. Fei (2015), Quantum Information Processing **14** (6), 2227.

Choi, M. (1975), Linear Algebra Appl. **10** (3), 285.

Christandl, M., and A. Winter (2005), IEEE Transactions on Information Theory **51** (9), 3159.

Coles, P. J. (2012), Physical Review A **85** (4), 042103.

Coles, P. J., R. Colbeck, L. Yu, and M. Zwolak (2012), Physical Review Letters **108** (21), 210405.

Coles, P. J., and F. Furrer (2015), Physics Letters A **379** (3), 105.

Coles, P. J., J. Kaniewski, and S. Wehner (2014), Nature Communications **5**, 5814.

Coles, P. J., and M. Piani (2014a), Physical Review A **89** (1), 010302.

Coles, P. J., and M. Piani (2014b), Physical Review A **89** (2), 022112.

Coles, P. J., L. Yu, V. Gheorghiu, and R. Griffiths (2011), Physical Review A **83** (6), 062338.

Cover, T. M., and J. A. Thomas (1991), *Elements of Information Theory* (Wiley).

Damgaard, I. B., S. Fehr, R. Renner, L. Salvail, and C. Schaffner (2007), in *Proc. CRYPTO 2007*, LNCS, Vol. 4622 (Springer) pp. 360–378.

Damgaard, I. B., S. Fehr, L. Salvail, and C. Schaffner (2008), SIAM Journal of Computing **37** (6), 1865.

Dammeier, L., R. Schwonnek, and R. F. Werner (2015), New Journal of Physics **17** (9), 093046.

Dankert, C., R. Cleve, J. Emerson, and E. Livine (2009), Physical Review A **80** (1), 012304.

Deutsch, D. (1983), Physical Review Letters **50** (9), 631.

Devetak, I., and A. Winter (2003), Physical Review A **68** (4), 042301.

Devetak, I., and A. Winter (2005), Proceedings of the Royal Society A **461** (2053), 207.

Dey, A., T. Pramanik, and A. S. Majumdar (2013), Physical Review A **87** (1), 012120.

Dietz, K. (2006), Journal of Physics A: Mathematical and Theoretical **39** (6), 1433.

DiVincenzo, D., M. Horodecki, D. Leung, J. Smolin, and B. Terhal (2004), Physical Review Letters **92** (6), 067902.

Dupuis, F., O. Fawzi, and S. Wehner (2015), IEEE Transactions on Information Theory **61** (2), 1093.

Dupuis, F., J. Florjanczyk, P. Hayden, and D. Leung (2013), Proceedings of the Royal Society A **469** (2159), 20130289.

Durt, T., B.-G. Englert, I. Bengtson, and K. Życzkowski (2010), International Journal of Quantum Information **08** (04), 535.

Dziembowski, S., and U. Maurer (2004), in *Proc. EUROCRYPT 2004*, LNCS, Vol. 3027 (Springer) pp. 126–137.

Eberle, T., V. Händchen, and R. Schnabel (2013), Optics Express **21** (9), 11546.

Einstein, A., B. Podolsky, and N. Rosen (1935), Physical Review **47** (10), 777.

Ekert, A. K. (1991), Physical Review Letters **67** (6), 661.

Englert, B.-G. (1996), Physical Review Letters **77** (11), 2154.

Englert, B. G., and J. a. Bergou (2000), Optics Communications **179** (1), 337.

Everett, H. (1957), Review of Modern Physics **29** (3), 454.

Fawzi, O., P. Hayden, and P. Sen (2011), in *Proc. ACM STOC 2011* (ACM Press, New York, NY) pp. 773–782.

Feng, J., Y.-Z. Zhang, M. D. Gould, and H. Fan (2013), Physics Letters B **726** (1-3), 527.

Frank, R. L., and E. H. Lieb (2012), Annals Henri Poincaré **13** (8), 1711.

Frank, R. L., and E. H. Lieb (2013a), Communications in Mathematical Physics **323** (2), 487.

Frank, R. L., and E. H. Lieb (2013b), Journal of Mathematical Physics **54** (12), 122201.

Friedland, S., V. Gheorghiu, and G. Gour (2013), Physical Review Letters **111** (23), 230401.

Furrer, F., J. Aberg, and R. Renner (2011), Communications in Mathematical Physics **306** (1), 165.

Furrer, F., M. Berta, M. Tomamichel, V. B. Scholz, and M. Christandl (2014), Journal of Mathematical Physics **55**, 122205.

Furrer, F., T. Franz, M. Berta, A. Leverrier, V. B. Scholz, M. Tomamichel, and R. F. Werner (2012), Physical Review Letters **109** (10), 100502.

Gavinsky, D., J. Kempe, I. Kerenidis, R. Raz, and R. de Wolf (2009), SIAM Journal of Computing **38** (5), 1695.

Gehring, T., V. Händchen, J. Duhme, F. Furrer, T. Franz, C. Pacher, R. F. Werner, and R. Schnabel (2015), Nature Communications **6**, 8795.

Ghirardi, G., L. Marinatto, and R. Romano (2003), Physics Letters A **317** (1-2), 32.

Gibbs, J. W. (1876), Transactions of the Connecticut Academy of Arts and Sciences **III**, 108.

Giovannetti, V., S. Lloyd, and L. Maccone (2011), Nature Photonics **5** (4), 222.

Gour, G., V. Narasimhachar, A. Grudka, M. Horodecki, W. Klobus, and J. Lodyga (2015), "The Conditional Uncertainty Principle," arXiv:1506.07124.

Grudka, A., M. Horodecki, P. Horodecki, R. Horodecki, W. Klobus, and L. Pankowski (2013), Physical Review A **88** (3), 032106.

Guha, S., P. Hayden, H. Krovi, S. Lloyd, C. Lupo, J. H.

Shapiro, M. Takeoka, and M. M. Wilde (2014), Physical Review X **4** (1), 011016.
Gühne, O., and G. Tóth (2009), Physics Reports **474** (1-6), 1.
Hall, M. (1995), Physical Review Letters **74** (17), 3307.
Hall, M. J. (1993), Journal of Modern Optics **40** (5), 809.
Hall, M. J. W. (1994), Physical Review A **49** (1), 42.
Hall, M. J. W. (1997), Physical Review A **55** (1), 100.
Hall, M. J. W. (1999), Physical Review A **59** (4), 2602.
Hall, M. J. W. (2004), Physical Review A **69** (5), 052113.
Hall, M. J. W. (2008), Journal of Physics A **41** (25), 255301.
Hall, M. J. W., D. W. Berry, M. Zwierz, and H. M. Wiseman (2012), Physical Review A **85** (4), 041802.
Hall, M. J. W., and H. M. Wiseman (2012), New Journal of Physics **14** (3), 033040.
Hasegawa, J., M. Hayashi, T. Hiroshima, and A. Tomita (2007), "Security Analysis of Decoy State Quantum Key Distribution Incorporating Finite Statistics," arXiv:0707.3541.
Hausladen, P., and W. K. Wootters (1994), Journal of Modern Optics **41** (12), 2385.
Hayden, P., D. Leung, P. W. Shor, and A. Winter (2004), Communications in Mathematical Physics **250** (2), 1.
Hayden, P., and J. Preskill (2007), Journal of High Energy Physics **2007** (09), 120.
Heisenberg, W. (1927), Zeitschrift für Physik **43** (3-4), 172.
Hiai, F., and D. Petz (1991), Communications in Mathematical Physics **143** (1), 99.
Hirschman, I. I. (1957), American Journal of Mathematics **79** (1), 152.
Holevo, A. S. (1973), Problems of Information Transmission **9** (3), 177.
Horodecki, M., J. Oppenheim, and A. Winter (2006), Communications in Mathematical Physics **269** (1), 107.
Horodecki, R., P. Horodecki, M. Horodecki, and K. Horodecki (2009), Review of Modern Physics **81** (2), 865.
Hu, M.-L., and H. Fan (2013), Physical Review A **88** (1), 014105.
Huang, Y. (2011), Physical Review A **83** (5), 052124.
Impagliazzo, R., L. A. Levin, and M. Luby (1989), in *Proc. ACM STOC 1989* (ACM Press) pp. 12–24.
Impagliazzo, R., and D. Zuckerman (1989), in *Proc. IEEE FOCS 1989*, pp. 248–253.
Ivanovic, I. D. (1992), Journal of Physics A: Mathematical and Theoretical **25** (7), L363.
Jaeger, G., A. Shimony, and L. Vaidman (1995), Physical Review A **51** (1), 54.
Jamiołkowski, A. (1972), Reports on Mathematical Physics **3** (4), 275.
Jia, L., Z. Tian, and J. Jing (2015), Annals of Physics **353**, 37.
Julsgaard, B., J. Sherson, J. I. Cirac, J. Fiurášek, and E. S. Polzik (2004), Nature **432** (7016), 482.
Kalev, A., and G. Gour (2014), New Journal of Physics **16** (5), 053038.
Kaniewski, J., M. Tomamichel, E. Hanggi, and S. Wehner (2013), IEEE Transactions on Information Theory **59** (7), 4687.
Kaniewski, J., M. Tomamichel, and S. Wehner (2014), Physical Review A **90** (1), 012332.
Kennard, E. H. (1927), Zeitschrift für Physik **44** (4-5), 326.
Kilian, J. (1988), in *Proc. ACM STOC 1988* (ACM Press, New York, NY) pp. 20–31.
Klappenecker, A., and M. Rotteler (2005), in *Proc. IEEE ISIT 2005* (IEEE) pp. 1740–1744.
Koashi, M. (2006), Journal of Physics: Conference Series **36** (1), 98.
König, R., and R. Renner (2011), IEEE Transactions on Information Theory **57** (7), 4760.
König, R., R. Renner, A. Bariska, and U. M. Maurer (2007), Physical Review Letters **98** (14), 140502.
König, R., R. Renner, and C. Schaffner (2009), IEEE Transactions on Information Theory **55** (9), 4337.
König, R., S. Wehner, and J. Wullschleger (2012), IEEE Transactions on Information Theory **58** (3), 1962.
Korzekwa, K., D. Jennings, and T. Rudolph (2014a), Physical Review A **89** (5), 052108.
Korzekwa, K., M. Lostaglio, D. Jennings, and T. Rudolph (2014b), Physical Review A **89** (4), 042122.
Kraus, K. (1987), Physical Review D **35** (10), 3070.
Krishna, M., and K. R. Parthasarathy (2002), Indian Journal of Statistics **64** (3), 842.
Kuznetsova, A. A. (2011), Theory of Probability and its Applications **55** (4), 709.
Larsen, U. (1990), Journal of Physics A: Mathematical and Theoretical **23** (7), 1041.
Li, C.-F., J.-S. Xu, X.-Y. Xu, K. Li, and G.-C. Guo (2011), Nature Physics **7** (10), 752.
Lieb, E. H., and M. B. Ruskai (1973), Journal of Mathematical Physics **14** (12), 1938.
Lindblad, G. (1975), Communications in Mathematical Physics **40** (2), 147.
Liu, S., L.-Z. Mu, and H. Fan (2015), Physical Review A **91** (4), 042133.
Liu, Y.-K. (2014), in *Proc. CRYPTO 2014*, LNCS, Vol. 8617, edited by J. A. Garay and R. Gennaro (Springer, Santa Barbara, CA) pp. 19–36.
Liu, Y.-K. (2015), in *Proc. EUROCRYPT 2015*, LNCS, Vol. 9057 (Springer) pp. 785–814.
Lo, H.-k. (1997), Physical Review A **56** (2), 1154.
Lo, H.-k., and H. F. Chau (1997), Physical Review Letters **78** (17), 3410.
Luo, S. L. (2005), Theoretical and Mathematical Physics **143** (2), 681.
Maassen, H., and J. Uffink (1988), Physical Review Letters **60** (12), 1103.
Maccone, L., and A. K. Pati (2014), Physical Review Letters **113** (26), 260401.
Mandayam, P., and S. Wehner (2011), Physical Review A **83** (2), 022329.
Mandayam, P., S. Wehner, and N. Balachandran (2010), Journal of Mathematical Physics **51** (8), 082201.
Marshall, A. W., I. Olkin, and B. C. Arnold (2011), *Inequalities: Theory of Majorization and Its Applications*, Springer Series in Statistics (Springer, New York).
Matthews, W., S. Wehner, and A. Winter (2009), Communications in Mathematical Physics **291** (3), 813.
Maurer, U. M. (1992), Journal of Cryptology **5** (1), 53.
Mayers, D. (1996), in *Proc. CRYPTO 1996*, LNCS, Vol. 1109 (Springer) pp. 343–357.
Mayers, D. (1997), Physical Review Letters **78** (17), 3414.
Mayers, D. (2001), Journal of the ACM **48** (3), 351.
Mclnnes, J. (1987), Technical Report, University of Toronto.
Miller, C. A., and Y. Shi (2014), in *Proc. ACM STOC 2014*, pp. 417–426.
Modi, K., A. Brodutch, H. Cable, T. Paterek, and V. Vedral (2012), Review of Modern Physics **84** (4), 1655.
Mosonyi, M., and T. Ogawa (2015), Communications in

Mathematical Physics **334** (3), 1617.
Müller-Lennert, M., F. Dupuis, O. Szehr, S. Fehr, and M. Tomamichel (2013), Journal of Mathematical Physics **54** (12), 122203.
Namiki, R., and Y. Tokunaga (2012), Physical Review Letters **108** (23), 230503.
Narasimhachar, V., A. Poostindouz, and G. Gour (2015), "The principle behind the Uncertainty Principle," arXiv:1505.02223.
von Neumann, J. (1932), *Mathematische Grundlagen der Quantenmechanik* (Springer).
Ng, N. H. Y., M. Berta, and S. Wehner (2012), Physical Review A **86** (4), 042315.
Ohya, M., and D. Petz (1993), *Quantum Entropy and Its Use* (Springer).
Ollivier, H., and W. H. Zurek (2001), Physical Review Letters **88** (1), 017901.
Oppenheim, J., and S. Wehner (2010), Science **330** (6007), 1072.
Ozawa, M. (2003), Physical Review A **67** (4), 042105.
Partovi, M. (1983), Physical Review Letters **50** (24), 1883.
Partovi, M. H. (2011), Physical Review A **84** (5), 052117.
Pati, A. K., M. M. Wilde, A. R. U. Devi, and A. K. Rajagopal (2012), Physical Review A **86** (4), 042105.
Portmann, C., and R. Renner (2014), "Cryptographic Security of Quantum Key Distribution," arXiv:1409.3525.
Prevedel, R., D. R. Hamel, R. Colbeck, K. Fisher, and K. J. Resch (2011), Nature Physics **7** (10), 757.
Puchała, Z., L. Rudnicki, K. Chabuda, M. Paraniak, and K. Życzkowski (2015), Physical Review A **92** (3), 032109.
Puchała, Z., L. Rudnicki, and K. Życzkowski (2013), Journal of Physics A: Mathematical and Theoretical **46** (27), 272002.
Rastegin, A. E. (2012), Quantum Information and Computation **12** (9), 0743.
Rastegin, A. E. (2013), European Physics Journal D **67** (12), 269.
Rastegin, A. E. (2014), Communications in Theoretical Physics **61** (3), 293.
Rastegin, A. E. (2015a), Quantum Information Processing **14** (2), 783.
Rastegin, A. E. (2015b), Open Systems & Information Dynamics **22** (01), 1550005.
Ren, L.-H., and H. Fan (2014), Physical Review A **90** (5), 052110.
Renes, J., and J.-C. Boileau (2009), Physical Review Letters **103** (2), 020402.
Renes, J. M., R. Blume-Kohout, A. J. Scott, and C. M. Caves (2004), Journal of Mathematical Physics **45** (6), 2171.
Renes, J. M., and J.-C. Boileau (2008), Physical Review A **78** (3), 032335.
Renes, J. M., and V. B. Scholz (2014), "Operationally-Motivated Uncertainty Relations for Joint Measurability and the Error-Disturbance Tradeoff," arXiv:1402.6711.
Renes, J. M., D. Sutter, F. Dupuis, and R. Renner (2015), IEEE Transactions on Information Theory **61** (11), 6395.
Renes, J. M., and M. M. Wilde (2014), IEEE Transactions on Information Theory **60** (6), 3090.
Renner, R. (2005), *Security of Quantum Key Distribution*, Ph.D. thesis (ETH Zurich).
Renner, R., and R. König (2005), in *Proc. TCC 2005*, LNCS, Vol. 3378 (Cambridge, MA) pp. 407–425.
Rényi, A. (1961), in *Proc. 4th Berkeley Symposium on Mathematical Statistics and Probability*, Vol. 1 (University of California Press, Berkeley, California, USA) pp. 547–561.
Ribeiro, J., and F. Grosshans (2015), .
del Rio, L., J. Aberg, R. Renner, O. Dahlsten, and V. Vedral (2011), Nature **474** (7349), 61.
Robertson, H. P. (1929), Physical Review **34** (1), 163.
Romera, E., and M. Calixto (2015), Journal of Physics: Condensed Matter **27** (17), 175003.
Rudnicki, L. (2011), Journal of Russian Laser Research **32** (4), 393.
Rudnicki, L. (2015), Physical Review A **91** (3), 032123.
Rudnicki, L., Z. Puchała, and K. Życzkowski (2014), Physical Review A **89** (5), 052115.
Rudnicki, L., S. P. Walborn, and F. Toscano (2012), Physical Review A **85** (4), 042115.
Rumin, M. (2011), Duke Mathematical Journal **160** (3), 567.
Rumin, M. (2012), Letters in Mathematical Physics **100** (3), 291.
Sánches-Ruiz, J. (1998), Physics Letters A **244** (4), 189.
Sánchez, J. (1993), Physics Letters A **173** (3), 233.
Sánchez-Ruiz, J. (1995), Physics Letters A **201** (2-3), 125.
Scarani, V., H. Bechmann-Pasquinucci, N. Cerf, M. Dusek, N. Lütkenhaus, and M. Peev (2009), Review of Modern Physics **81** (3), 1301.
Schaffner, C. (2007), *Cryptography in the Bounded-Quantum-Storage Model*, Phd thesis (University of Aarhus).
Schneeloch, J., C. J. Broadbent, and J. C. Howell (2014), Physical Review A **90** (6), 062119.
Schneeloch, J., C. J. Broadbent, S. P. Walborn, E. G. Cavalcanti, and J. C. Howell (2013), Physical Review A **87** (6), 062103.
Schrödinger, E. (1930), Proceedings of the Prussian Academy of Sciences **XIX**, 296.
Schumacher, B., and M. A. Nielsen (1996), Physical Review A **54** (4), 2629.
Shannon, C. (1948), Bell System Technical Journal **27**, 379.
Shor, P. W., and J. Preskill (2000), Physical Review Letters **85** (2), 441.
Slepian, D., and H. O. Pollak (1961), Bell System Technical Journal **40** (1), 43.
Stinespring, W. F. (1955), Proceedings of the Americal Mathematical Society **6**, 211.
Tomamichel, M. (2012), *A Framework for Non-Asymptotic Quantum Information Theory*, Ph.D. thesis (ETH Zurich).
Tomamichel, M. (2016), *Quantum Information Processing with Finite Resources — Mathematical Foundations*, SpringerBriefs in Mathematical Physics, Vol. 5 (Springer International Publishing).
Tomamichel, M., M. Berta, and M. Hayashi (2014), Journal of Mathematical Physics **55** (8), 082206.
Tomamichel, M., R. Colbeck, and R. Renner (2009), IEEE Transactions on Information Theory **55** (12), 5840.
Tomamichel, M., R. Colbeck, and R. Renner (2010), IEEE Transactions on Information Theory **56** (9), 4674.
Tomamichel, M., S. Fehr, J. Kaniewski, and S. Wehner (2013), New Journal of Physics **15** (10), 103002.
Tomamichel, M., and E. Hänggi (2013), Journal of Physics A: Mathematical and Theoretical **46** (5), 055301.
Tomamichel, M., C. C. W. Lim, N. Gisin, and R. Renner (2012), Nature Communications **3**, 634.
Tomamichel, M., and R. Renner (2011), Physical Review Letters **106** (11), 110506.
Tsallis, C. (1988), Journal of Statistical Physics **52** (1-2), 479.
Uffink, J. (1990), *Measures of Uncertainty and the Uncertainty Principle*, Ph.D. thesis (R.U. Utrecht).

Uhlmann, A. (1977), Communications in Mathematical Physics **54** (1), 21.

Umegaki, H. (1962), Kodai Math. Sem. Rep. **14**, 59.

Unruh, D. (2010), in *Proc. EUROCRYPT 2010*, LNCS, Vol. 6110 (Springer) pp. 486–505.

Vadhan, S. P. (2012), *Pseudorandomness*, Foundations and Trends in Theoretical Computer Science, Vol. 7 (Now Publishers).

Vallone, G., D. G. Marangon, M. Tomasin, and P. Villoresi (2014), Physical Review A **90** (5), 052327.

Ver Steeg, G., and S. Wehner (2009), Quantum Information and Computation **9** (9&10), 0801.

de Vicente, J., and J. Sánchez-Ruiz (2008), Physical Review A **77** (4), 042110.

Walborn, S. P., A. Salles, R. M. Gomes, F. Toscano, and P. H. Souto Ribeiro (2011), Physical Review Letters **106** (13), 130402.

Weedbrook, C., S. Pirandola, R. García-Patrón, N. Cerf, T. Ralph, J. Shapiro, and S. Lloyd (2012), Review of Modern Physics **84** (2), 621.

Wehner, S., C. Schaffner, and B. M. Terhal (2008), Physical Review Letters **100** (22), 220502.

Wehner, S., and A. Winter (2008), Journal of Mathematical Physics **49** (6), 062105.

Wehner, S., and A. Winter (2010), New Journal of Physics **12** (2), 025009.

Wehrl, A. (1978), Review of Modern Physics **50** (2), 221.

Weyl, H. (1928), *Gruppentheorie und Quantenmechanik* (Hirzel, Leipzig).

Wilde, M. M., A. Winter, and D. Yang (2014), Communications in Mathematical Physics **331** (2), 593.

Winter, A., and D. Yang (2015), "Operational Resource Theory of Coherence," arXiv:1506.07975.

Wiseman, H. M., S. J. Jones, and A. C. Doherty (2007), Physical Review Letters **98** (14), 140402.

Wootters, W., and W. H. Zurek (1979), Physical Review D **19** (2), 473.

Wootters, W. K., and B. D. Fields (1989), Annals of Physics **191** (2), 363.

Wootters, W. K., and D. M. Sussman (2007), in *Proc. QCMC 2007*, p. 269, arXiv:0704.1277.

Wootters, W. K., and W. H. Zurek (1982), Nature **299** (5886), 802.

Wu, S., S. Yu, and K. Molmer (2009), Physical Review A **79** (2), 022104.

Zozor, S., G. M. Bosyk, and M. Portesi (2014), Journal of Physics A: Mathematical and Theoretical **47** (49), 495302.

Zurek, W. H. (2003), Review of Modern Physics **75** (3), 715.

Życzkowski, K., and I. Bengtsson (2004), Open Systems & Information Dynamics **11** (01), 3.