One-parameter class of uncertainty relations based on entropy power

Petr Jizba,1,2,* Yue Ma,3 Anthony Hayes,4 and Jacob A. Dunningham4,1
1FNSPE, Czech Technical University in Prague, Břehová 7, 115 19 Praha 1, Czech Republic
2ITP, Freie Universität Berlin, Arnimallee 14 D-14195 Berlin, Germany
3Department of Physics, Tsinghua University, Beijing 100084, People’s Republic of China
4Department of Physics and Astronomy, University of Sussex, Brighton BN1 9QH, United Kingdom

(Received 23 March 2016; published 29 June 2016)

We use the concept of entropy power to derive a one-parameter class of information-theoretic uncertainty relations for pairs of conjugate observables in an infinite-dimensional Hilbert space. This class constitutes an infinite tower of higher-order statistics uncertainty relations, which allows one in principle to determine the shape of the underlying information-distribution function by measuring the relevant entropy powers. We illustrate the capability of this class by discussing two examples: superpositions of vacuum and squeezed states and the Cauchy-type heavy-tailed wave function.

DOI: 10.1103/PhysRevE.93.060104

Introduction. In 1948, Shannon laid down the foundations of modern information theory [1]. He was instrumental in pointing out that, in contrast with discrete signals or messages where information is quantified by (Shannon’s) entropy, the case with continuous variables is less satisfactory. The continuous version of Shannon’s entropy (SE)—the so-called differential entropy—may take negative values [1,2] and so does not have the same status as its discrete-variable counterpart. To solve a range of communication-theoretic problems related to continuous cases Shannon shifted the emphasis from the differential entropy to another object—entropy power (EP). The EP represents the variance of a would-be Gaussian random variable with the same differential entropy as the random variable under investigation. EP was used by Shannon [1] to bound the capacity of non-Gaussian additive noise channels. Subsequent developments in information theory confirmed the central role of the EP [3–5]. On the mathematical side the EP proved to be critical in proving a strong version of the central limit theorem with convergence in relative entropy [6,7].

Information theory now extends far beyond the realm of communications and the same principles and concepts can be employed in applications that include statistical physics, biological science, and quantum mechanics [8]. In this Rapid Communication we focus on the application of the EP to quantum-mechanical uncertainty relations (URs). In essence, quantum-mechanical URs place fundamental limits on the accuracy with which one is able to know the values of different physical quantities. In the 1920s, Kennard, and independently, Robertson and Schrödinger reformulated the original Heisenberg’s UR in terms of variances of the observables [9–11]. In 1959, Stam [12] conjectured that the EP could be used to obtain Heisenberg’s UR. This conjecture was bolstered in [13] by showing that the usual Schrödinger-Robertson variance-based URs (VURs) [10,11] can be derived from entropic URs. VURs are useful and widely applied but have two major restrictions: First, the product of the conjugate variances is a single number and so can only ever give partial information about the underlying states; secondly, variances are only useful concepts for well-behaved bell-like distributions. For heavy-tailed or multipeaked distributions, the variances can be large or even infinite, making VURs ill suited or even useless.

Here we show that Stam’s UR (and VUR) is just a member of a one-parameter class of EP-based inequalities, all of which stem from yet another important information measure, namely, the Rényi entropy (RE) [14,15] and its continuous counterparts, differential RE and Rényi entropy power (REP). We prove that this class constitutes an infinite tower of higher-order cumulant URs, which allows one in principle to reconstruct the underlying information-distribution function in a process akin to quantum state tomography [16] using EPs in place of the usual measurements. In this respect, the strategy is not to optimize parameters in the class of URs (e.g., to find a best bound), but instead to identify and measure as many EPs (associated with a given quantum state) as possible. We illustrate this point with two examples of interest.

Entropy power. Let $\mathcal{X}$ be a random vector in $\mathbb{R}^D$ with the probability density function (PDF), $\mathcal{F}$. The differential entropy $H(\mathcal{X})$ of $\mathcal{X}$ is defined as [1]

$$H(\mathcal{X}) = -\int_{\mathbb{R}^D} \mathcal{F}(x) \log_2 \mathcal{F}(x) dx.$$ (1)

The discrete version of (1) is nothing but the SE [1]. Strictly the form shown in (1) is not a proper entropy but rather an information gain [2,14]. The entropy power $N(\mathcal{X})$ of $\mathcal{X}$ is the unique number such that [1,17]

$$H(\mathcal{X}) = H(\sqrt{N(\mathcal{X})} Z_G),$$ (2)

with $Z_G$ representing a Gaussian random vector with zero mean and unit covariance matrix. In the case when the Shannon differential entropy is measured in nats (natural units) the entropy power takes the form [1]

$$N(\mathcal{X}) = \frac{1}{2\pi e} \exp \left( \frac{2}{D} H(\mathcal{X}) \right).$$ (3)

Correspondingly, the differential Rényi entropy $I_{\rho}(\mathcal{X})$ of $\mathcal{X}$ is defined as [2,14]

$$I_{\rho}(\mathcal{X}) = \frac{1}{(1-\rho)} \log_2 \left( \int_{\mathbb{R}^D} dx \mathcal{F}^\rho(x) \right).$$ (4)
where the index $p \in \mathbb{R}^+$. With the help of L'Hôpital’s rule one can check that for $p \to 1$ one has $I_p(\chi) \to H(\chi)$. Similarly to $H$, $I_p$ is also additive for independent events [2]. In analogy with the case of Shannon entropy discussed above, the $p$th Rényi entropy power $N_p(\chi)$ is defined as the solution of the equation

$$ I_p(\chi) = I_p(\sqrt{N_p(\chi)} Z_G), $$

where $Z_G$ represents a Gaussian random vector with zero mean and unit covariance matrix.

This type of expression was studied in [18,19] where it was shown that the only class of solution of (5) is

$$ N_p(\chi) = \frac{1}{2\pi} p^{-\rho/p} \exp\left(\frac{2}{D} I_p(\chi)\right), $$

with $1/p + 1/p' = 1$ and $p \in \mathbb{R}^+$. In addition, when $p \to 1+$ one has $N_p(\chi) \to N(\chi)$. For simplicity we have taken nats as units of information. In passing, we may observe that from (6) it follows that $N_p(\sigma Z_G) = \sigma^2$, i.e., for Gaussian processes the EP is simply the variance $\sigma^2$. In the case where $Z_G$ represents a random Gaussian vector of zero mean and covariance matrix $K_{ij}$, then $N_p(Z_G) = |\det(K_{ij})|^{1/D} = |K|^{1/D}$. Importantly, since the REs are in principle measurable [20,21], the associated REPs are experimentally accessible. For some recent applications of the REs in quantum theory see, e.g., [22–24].

**Entropy power uncertainty relations.** We start with the theorem of Beckner and Babenko [25,26].

**Beckner-Babenko Theorem.** Let

$$ f^{(2)}(x) = \int_{\mathbb{R}^D} e^{2\pi i xy} f^{(1)}(y) dy, $$

then for $p \in [1,2]$,

$$ |(p^{D/2})^{1/p} f^{(2)}(p)| \leq |p^{D/2})^{1/p} f^{(1)}|_p, $$

where $p$ and $p'$ are the Hölder conjugates and

$$ |F|_p \equiv \left(\int_{\mathbb{R}^D} |F(y)|^p dy\right)^{1/p}, $$

for any $F \in L^p(\mathbb{R}^D)$. Of course, the role of $f^{(1)}$ and $f^{(2)}$ may be interchanged in the inequality (7). An elementary proof can be found, e.g., in [18]. Inequality (7) is saturated only for Gaussian functions [26,27].

Anticipating quantum-mechanical applications we define $\sqrt{F(y)} \equiv |f(y)|$. After some simple algebra we recast (7) in the form [18]

$$ \left(\int_{\mathbb{R}^D} |F^{(2)}(y)|^{1+\tau} dy\right)^{1/(1+\tau)} \leq \left(\int_{\mathbb{R}^D} |F^{(1)}(y)|^{1+\tau} dy\right)^{1/(1+\tau)}. $$

Here, $r = p/D - 1$ and $t = p'/2 - 1$. Because $1/p + 1/p' = 1$ we have the constraint $-r/(2r + 1)$. Since $p \in [1,2]$ one has $r \in [-1/2,0]$ and $t \in [0,\infty)$. Taking the negative binary logarithm of both sides of (9), we obtain

$$ I_{1+r}(F^{(2)}) + I_{1+r}(F^{(1)}) \geq \frac{1}{r} \log_2[2(1+t)]^{D/2} + \frac{1}{t} \log_2[2(1+t)]^{D/2}. $$

In the limit $t \to 0+$ and $r \to 0-$ this reduces to

$$ \mathcal{H}(F^{(2)}) + \mathcal{H}(F^{(1)}) \geq \log_2 \left(\frac{e}{2}\right)^D, $$

which is just the classical Hirschman conjecture for Shannon’s differential entropies [13,28]. However, the semidefiniteness of $I_p(\cdots)$ makes the URs (10) impractical. In terms of REPs we can rewrite (10) as

$$ N_{1+r}(F^{(2)}) + N_{1+r}(F^{(1)}) \geq N_{p/2}(\chi)N_{q/2}(\chi) \geq \frac{1}{16\pi^2}. $$

The equality follows from the saturation of the inequality (7) by Gaussian functions.

By assuming that a PDF has a finite covariance matrix $(K_{ij})$ then important inequalities hold, namely,

$$ N(\chi) \leq |K|^{1/D} = \sigma^2, $$

with equality in the first inequality if and only if $\chi$ is a Gaussian vector, and in the second if and only if $\chi$ has a covariance matrix that is proportional to the identity matrix. The proof of (14) is based on the non-negativity of the Kullback-Leibler divergence and can be found, e.g., in [29,30]. Inequality (14) immediately gives

$$ \sigma^2 \geq \log_2 Z \equiv \log_2 \left[\frac{1}{2\pi e} \right]^{D/2}, $$

which saturates only for Gaussian (respective white) random vectors $\chi$ and $\gamma$. Note, that when $(K_{ij})$ and $(K_{ij})$ exist then (15) automatically implies the conventional Robertson-Schrödinger VUR. Since the VUR is implied by the Shannon EPUR alone, a natural question arises; in what sense is the general set of inequalities (12) more informative than the special case $r = t = 0$?

**Reconstruction theorem.** To aid our intuition and, furthermore, to show the conceptual underpinning for Rényi-entropy-power-based uncertainty relations (REPURs) (12) we first note that the differential RE can be written as (in units of bits$^a$).

$$ \frac{1}{(1-p)} \log_2 E[2^{1-p}r(\chi)]. $$

Here $i(\chi)$ is the information in $\chi$ with respect to the PDF $f(\chi)$. From (16), the differential RE can be viewed as a reparametrized version of the cumulant generating function of the information random variable $i(\chi)$. The ensuing cumulant expansion is

$$ pI_{1-p}(\chi) = \log_2 e \sum_{n=1}^{\infty} \frac{\kappa_n(\chi)}{n!} \left(\frac{p}{\log_2 e}\right)^n, $$

where $\kappa_n(\chi) \equiv \kappa_n(i(\chi))$ denotes the $n$th cumulant of $i(\chi)$ (in units of bits$^a$). From (17) it follows that REPs can be
written in terms of $\kappa_n$’s. In fact, $N_p$’s of order $p > 0$ uniquely determine the underlying information PDF (for the proof see Supplemental Material [31]). So, the REPURs of different orders provide additional structural constraints between $\mathcal{F}^{(1)}$ and $\mathcal{F}^{(2)}$ which cannot be seen with the VUR or Shannon entropy UR alone. In this connection we list some further salient results [31]:

(a) Only Gaussian PDFs saturate all REPURs. REPURs with $r = -1/2$ can be saturated with a wider class of PDFs. (b) When $\mathcal{F}^{(x)}$ is close to (or equimeasurable with) a Gaussian PDF then only $N_p$’s with $p$’s in a neighborhood of 1 are needed. The closer the shape is to the Gaussian PDF, the more accentuated. So, when the accentuated parts in $p < 1$ are Gaussian, though both $\mathcal{F}^{(1)}$ and $\mathcal{F}^{(2)}$ might be non-Gaussian.

**REPR in quantum mechanics.** Let us consider state vectors that are Fourier transform duals—the most prominent example being the configuration and momentum space wave functions. In such a case there is a reciprocal relation between $\psi(x)$ and $\hat{\psi}(p)$, namely,

$$\psi(x) = \int_{\mathbb{R}^d} e^{ipx/\hbar} \hat{\psi}(p) \frac{dp}{(2\pi \hbar)^{D/2}}. \tag{18}$$

The Riesz-Fischer equality [32] guarantees mutual normalization $\|\psi\|_2 = \|\hat{\psi}\|_2 = 1$. Let us define

$$f^{(2)}(x) = (2\pi \hbar)^{D/4} \psi(\sqrt{2\pi \hbar} x),$$

$$f^{(1)}(p) = (2\pi \hbar)^{D/4} \hat{\psi}(\sqrt{2\pi \hbar} p). \tag{19}$$

The factor $(2\pi \hbar)^{D/4}$ ensures that the new functions are normalized (in sense of $\mathcal{L}_2$) to unity. With these we have the same structure of the Fourier transform as in the Beckner-Babenko theorem. Consequently we can write the associated RE-based URs (10) in the form

$$\mathcal{I}_{1+r}(|\psi|^2) + \mathcal{I}_{1+r}(|\hat{\psi}|^2) \geq \frac{1}{r} \log_2 \left( \frac{1 + r}{\pi \hbar} \right)^{D/2} + \frac{1}{t} \log_2 \left( \frac{1 + t}{\pi \hbar} \right)^{D/2}, \tag{20}$$

where we have made use of the identity

$$\mathcal{I}_p(|f^{(1)}|^2) = \mathcal{I}_p(|\hat{\psi}|^2) = \frac{D}{2} \log_2(2\pi \hbar) \tag{21}$$

(and similarly for $f^{(2)}$). In terms of the REP we can recast (20) into the form [cf. Eq. (12)]

$$N_{1+r}(|\psi|^2) N_{1+r}(|\hat{\psi}|^2) \geq \frac{\hbar^2}{4}. \tag{22}$$

This looks similar to the Robertson-Schrödinger VUR, but is now a family of relations parametrized by $t$ (or equivalently $r$) each having the same universal lower bound $\hbar^2/4$. It should be noted that the familiar VUR follows directly from Shannon’s entropy power UR alone since [cf. Eq. (15)]

$$\sigma_x^2 \sigma_p^2 \geq N_1(|\psi|^2) N_1(|\hat{\psi}|^2) \geq \frac{\hbar^2}{4}. \tag{23}$$

In the special case of Gaussian PDFs, the whole family reduces to the single familiar coherent-state VUR

$$\sigma_x^2 \sigma_p^2 = N_{1+r}(|\psi_G|^2) N_{1+r}(|\hat{\psi}_G|^2) = \frac{\hbar^2}{4}. \tag{24}$$

**Applications in quantum mechanics.** As a first example we consider an optical state that is pertinent to quantum metrology [33]. It consists of a superposition of a vacuum $|0\rangle$ and a squeezed vacuum $|z_\zeta\rangle$ which has the form $|\psi_\zeta\rangle = N |0\rangle + |z_\zeta\rangle$, with $N = 1/\sqrt{2+2\cosh \zeta}$, and

$$|z_\zeta\rangle = \sum_{m=0}^\infty (-1)^m \frac{\sqrt{(2m)!}}{2^m m!} \left[ \frac{\tanh \zeta}{\sqrt{\cosh \zeta}} \right]^{2m} |2m\rangle, \tag{25}$$

where $|2m\rangle$ are even-number energy eigenstates and $\zeta \in \mathbb{R}$ is the squeezing parameter. If we rewrite $|\psi_\zeta\rangle$ in the basis of the eigenstates of the position and momentum quadrature operators

$$\hat{X} = \sqrt{\frac{\hbar}{2\omega}} (\hat{a} + \hat{a}^\dagger), \quad \hat{P} = -i \sqrt{\frac{\hbar \omega}{2}} (\hat{a} - \hat{a}^\dagger), \tag{26}$$

($\omega$ is the optical frequency and $\hat{a}$ and $\hat{a}^\dagger$ are, respectively, the photon annihilation and creation operators), we get for the PDFs (apart from normalization $N^2$)

$$|\psi_\zeta|^2 = \sqrt{\frac{\omega}{\pi \hbar}} \exp \left( -\frac{\omega x^2}{2\hbar} + e^{\zeta/2} \exp \left( -\frac{\omega e^{2\zeta} x^2}{2\hbar} \right) \right)^2, \tag{27}$$

$$|\hat{\psi}_\zeta|^2 = \frac{1}{\sqrt{\pi \hbar \omega}} \exp \left( -\frac{p^2}{2\hbar \omega} + e^{-\zeta/2} \exp \left( -\frac{e^{-2\zeta} p^2}{2\hbar \omega} \right) \right)^2. \tag{27}$$

These can be used to calculate the product $N_{1+r}(x) N_{1+r}(p)$ for different values of $r$. The result is depicted in Fig. 1 for three different values of the squeezing parameter. What we find is that the lower bound $\hbar^2/4$ is saturated for both $N_{1/2}(x) N_{1/2}(p)$ and $N_{1/2}(x) N_{1/2}(p)$ regardless of the squeezing (in Fig. 1 these correspond to $r = -1/2$ and $r \to \infty$, respectively). From our foregoing analysis of REPURs this is easy to understand because the infinite and half indices of the EPs focus on the peak and tails of the PDF, respectively, and from (27) we see that both the $x$ and $p$ PDFs are Gaussian in the tails as well as at the peaks (i.e., at $x = p = 0$). A REPUR is saturated only when the RE-distributed sectors in both dual PDFs are Gaussian [31]. On the other hand, it is also clear that both PDFs (27) as a whole are highly non-Gaussian. We would therefore not expect REPURs with different indices to saturate the bound. This is clearly illustrated in Fig. 1. In passing, we note that for any $\zeta \neq 0$ the Shannon entropy power UR is the furthest from saturating the bound, and so is the least informative of all the family of REPURs.
By way of comparison, we can also calculate the VUR for the state $|\psi_1\rangle$. The variances involved are

$$\langle (\Delta X)^2 \rangle_1 = N^2 \hbar \omega \left[ \frac{1}{2} (1 + e^{-2\xi}) + \sqrt{\text{sech}(1 + \tanh \xi)} \right],$$

$$\langle (\Delta P)^2 \rangle_1 = N^2 \hbar \omega \left[ \frac{1}{2} (1 + e^{2\xi}) + \sqrt{\text{sech}(1 + \tanh \xi)} \right].$$

For $\xi = 0$, we have $\langle (\Delta X)^2 \rangle_0 = \langle (\Delta P)^2 \rangle_0 = \hbar^2/4$, i.e., the VUR is saturated. This is no surprise because, in this case, the vacuum $|0\rangle$ is the usual (Glauber) coherent state. However, as the squeezing parameter $\xi$ is increased the product blows up rapidly, which makes the VUR uninformative. So the set of REPURs outperform both the Shannon EPUR and the VUR by providing more information on the structural features of $|\psi_1\rangle$ via the related PDFs (e.g., Gaussian peaks and tails in $p$-$x$ quadratures). A similar type of behavior can be also seen in a particular class of Schrödinger cat states represented by two superposed Glauber coherent states with the variable amplitude parameter [18]. In the aforesaid case the Fourier transform duals were chosen to be two orthogonal phase quadratures $|x_0, x_{1/2}\rangle$. Specifically for $r = -1/2$ and $r \rightarrow \infty$ it was observed that the entropic inequality (10) (and hence also the associated REPUR) were saturated for the amplitude parameter $\beta < 1/2$, which according to [31] implies Gaussianity of the respective tails and peaks in state PDFs. Since the REPUR is not saturated for $\beta \geq 1/2$ either peaks or tails cannot be Gaussian. Closer analysis indeed revealed that the state PDFs for $\beta \geq 1/2$ start to develop two separated peaks corresponding to the separation of two overlapping Gaussian wave packets. In addition, for any $r$ the REPURs are for large $\beta$ independent of the value of $\beta$. This is a consequence of two facts: (a) for large $\beta$ the two Gaussian wave packets no longer overlap and (b) REPUs are immune to piecewise rearrangements of the PDF [18,31].

We note that the conventional VUR does not pose any restriction on the variance of the observable whose conjugate observable has a PDF with infinite covariance matrix. So, such a state is maximally uncertain. In contrast to this, the set of related REPURs brings considerably more information about the structure of these states. To illustrate this we discuss in our second example a power-law tail wave packet (PLTP). PTLPWs are archetypal examples of quantum states with anomalous (scaling) behavior during their temporal evolution [34]. For definiteness we will consider the PLTP of the form

$$\psi(x) = \sqrt{\frac{\gamma}{\pi}} \left( \frac{1}{\gamma^2 + (x - m)^2} \right)^{\frac{1}{2}},$$

which entails the Cauchy PDF with a scale parameter $\gamma$ and median $m$. The Fourier transform reads

$$\hat{\psi}(p) = e^{-imp/\hbar} \left[ \frac{2\gamma}{\pi} \frac{1}{\sqrt{\pi^2 \hbar K_0(|\gamma|p|/\hbar)}} \right]$$

($K_0$ is the modified Bessel function). With these results we can immediately write two representative REPURs:

$$N_1(|\hat{\psi}|^2)N_1(|\psi|^2) = 0.0052\hbar^2\pi^4 > \hbar^2/4,$$

$$N_{1/2}(|\hat{\psi}|^2)N_{\infty}(|\psi|^2) = \hbar^2/4.$$  

Note also that $\langle (\Delta P)^2 \rangle_\psi = \hbar^2/16\pi^2$ and $\langle (\Delta X)^2 \rangle_\psi \rightarrow \infty$ (the latter behavior is symptomatic of many PTLPWs), and so the Schrödinger-Robertson’s VUR is completely uninformative. What can we conclude from (30) and (31)? First, the REPUR (31) is saturated. This implies that the peak part of $|\psi|^2$ and the tail part of $|\hat{\psi}|^2$ are Gaussian (as can be directly checked). Shannon’s EPUR (30) implies: (a) the involved PDFs are not Gaussian, (b) in contrast to other REPURs it quantifies only shape structures of PDFs but is $\gamma$ insensitive [31], and (c) from (11) [cf. also (21)] the lower bound of Hirschman’s UR is $\log_2(\pi \hbar e)$ while (30) gives $\log_2(\pi \hbar e) + 0.5141$, so one could still gain 0.5141 bits of information should the system be prepared in a Gaussian state. Finally, we note that $N_{1/2}(|\hat{\psi}|^2) = 0$ and $N_{1/2}(|\psi|^2) \rightarrow \infty$, hence the related REPUR is indeterminate (in fact, regularization dependent). This behavior is easy to understand. For a strongly leptokurtic PDF (such as $|\psi|^2$) $N_{1/2}$ accentuates the very flat power-law tails of $|\psi|^2$, and hence $N_{1/2}$ represents the variance of a very flat (almost equiprobable) Gaussian PDF. Similarly, $N_{\infty}$ accentuates only the peak part of $|\hat{\psi}|^2$ that is sharply (almost $\delta$ function) peaked, and so $N_{\infty}$ represents the variance of the Gaussian PDF with zero spread. Let us finally mention that in [31] it is shown how to deduce from REPUs the scaling characteristics for Lévy stable and Laplacian PTLPWs.

**Conclusions.** In this Rapid Communication we have formulated a one-parameter class of Rényi-entropy-power-based URs for pairs of observables in an infinite-dimensional Hilbert space. The tower of inequalities obtained possess a clear advantage over the single VUR by revealing the finer structure of the underlying PDFs further to their standard deviations. This was demonstrated on two relevant quantum-mechanical examples and mathematically substantiated via the
reconstruction theorem. We have also established a formal link between the Robertson-Schrödinger VUR and Shannon-Hirschman UR and highlighted the limited scope of the VUR. Notably, we have shown that the Robertson-Schrödinger VUR is a simple consequence of the REPUR with the index \( r = 0 \) while other REPURs in the class set fundamental (irreducible) limits on higher-order cumulants in conjugate information PDFs.

Acknowledgments. P.J. was supported by the GAČR Grant No. GA14-07983S. J.D. acknowledges support from DSTL and the UK EPSRC through the NQIT Quantum Technology Hub (Grant No. EP/M013243/1).