This is a repository copy of *Sharp uncertainty relations for number and angle*.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/id/eprint/129361/

Version: Accepted Version

Sharp uncertainty relations for number and angle

Paul Busch,[1, a)] Jukka Kiukas,[2, b)] and R.F. Werner[3, c)]

[1)]*Department of Mathematics, University of York, York,*
United Kingdom

[2)]*Department of Mathematics, Aberystwyth University, SY23 3BZ Aberystwyth,*
United Kingdom

[3)]*Institut für Theoretische Physik, Leibniz Universität, Hannover,*
Germany

We study uncertainty relations for pairs of conjugate variables like number and angle, of which one takes integer values and the other takes values on the unit circle. The translation symmetry of the problem in either variable implies that measurement uncertainty and preparation uncertainty coincide quantitatively, and the bounds depend only on the choice of two metrics used to quantify the difference of number and angle outputs, respectively. For each type of observable we discuss two natural choices of metric, and discuss the resulting optimal bounds with both numerical and analytic methods. We also develop some simple and explicit (albeit not sharp) lower bounds, using an apparently new method for obtaining certified lower bounds to ground state problems.

[a)]Electronic mail: paul.busch@york.ac.uk

[b)]Electronic mail: jek20@aber.ac.uk

[c)]Electronic mail: reinhard.werner@itp.uni-hannover.de

arXiv:1604.00566v2 [quant-ph] 30 Mar 2018

I. INTRODUCTION

The study of uncertainty relations has experienced a major boost in recent years. As more and more experiments reach quantum limited accuracy, sharp quantitative uncertainty and error bounds become more relevant. It has also become evident that the subject of quantum uncertainty cannot be reduced to the classic standard uncertainty relation that was first made rigorous by Kennard[1], Robertson[2] and others and is now found in every textbook. The purpose of this paper is to present a new case study that exemplifies the three main directions of generalization currently being pursued. The example at hand is given by the number-angle pair of observables, where by "number" we understand an observable whose spectrum is the set of all integers.

The first extension of the uncertainty principle concerns the set of scenarios to which quantitative uncertainty relations apply. The Kennard relation is a "preparation uncertainty relation", i.e., a quantitative expression of the observation that there is no state preparation for which the distributions of two observables under consideration are both sharp. This relation can be tested, in principle, by separate runs of precise measurements of the two observables, performed on ensembles of systems in the same state. In contrast to this, one may consider attempted joint measurements of noncommuting observables, as already intuitively envisaged by Heisenberg[3]; one finds that such joint measurements are constrained by "measurement uncertainty relations"[4–6] which describe the unavoidable error bounds.

The number-angle pair also highlights the need to consider uncertainty measures other than standard deviations. The second direction of generalization to be considered is thus in the concrete mathematical expressions measuring the "sharpness" of distributions, or the error of an approximate measurement. We will consider two alternative types of uncertainty measures and associated error measures for each of the two observables concerned.

Finally, the third direction of generalization concerns the forms of preparation and measurement uncertainty relations that are applicable to arbitrary pairs of observables. The Robertson relation involving the expectations of the commutator in the lower bound fails to provide preparation uncertainty relation in the above sense because the only state-independent lower bound one can get from it is zero. Nevertheless, non-trivial state-independent bounds usually do exist. Uncertainty is really a ubiquitous phenomenon, joint measurability or simultaneous sharp preparability are the exceptions rather than the rule.

Accordingly, the problem of establishing tight uncertainty relations for pairs of observables amounts to the task of establishing their (preparation or measurement) *uncertainty regions*, defined as the set of pairs of uncertainty values in all states (for preparation uncertainty) and the set of pairs of error values in all possible joint measurements (for measurement uncertainty), respectively.

It is a pleasing property of the case at hand—number and angle—that an essentially complete treatment can be given. This is due to the phase space symmetry which implies, exactly as for standard position and momentum[7], that "metric distance" and "calibration distance" for the error assessment of observables satisfy the same relations and also that they are quantitatively the same as corresponding preparation uncertainty relations. In this paper we deduce the ensuing measurement uncertainty relations for number and angle.

Our paper is organized according to the methods employed. In Section II we review the conceptualization of the preparation uncertainty measures and associated error measures to be used throughout the paper. This is followed by an overview of our main results (Section III). In Section IV we show how the phase space symmetry can be exploited to find optimal joint measurements among the covariant phase space observables. Here we show the identity of preparation uncertainty regions and measurement uncertainty (or error) regions in the case at hand, introducing calibration uncertainty regions as a mediating construct. The lower boundary of the uncertainty regions is characterized as a ground state problem. Numerical estimates of the optimal tradeoff curves are shown in Section V. Next, Sections VI and VII give determinations of the exact ground states and lower boundaries for the various combinations of deviation measures, and we show how existing uncertainty relations for number and phase can be reproduced or strengthened using our systematic approach. We conclude with an outlook in Section VIII.

II. CONCEPTUAL UNCERTAINTY BASICS

A. Number and angle observables

The complementarity between number and angle appears in physics in various guises. Essentially, this is for any parameter with a natural periodicity. Geometric angles are one case, with the complementary variable given by a component of angular momentum. Another

important case is quantum optical phase, which is complementary to an harmonic oscillator Hamiltonian. At least for preparation uncertainty this is no difficulty, since a relation valid for all states also holds for states supported on the subspace of positive integers. One does get additional or sharper relations from building in this constraint, however. A further important field of applications is quasi-momentum with values in the Brillouin-Zone for a lattice system (of which we only consider the one-dimensional case here). This is also related to form factors arising in the discussion of diffraction patterns and fringe contrast from periodic gratings[8].

The literature on angle-angular momentum uncertainty is almost exclusively concerned with the preparation scenario, although the lack of an error disturbance relation has been noted[9]. In the preparation case a major obstacle was the Kennard and Robertson[2] relation and their role as a model how uncertainty relations should be set up. There is nothing wrong with an observable with outcomes on a circle. But much work was wasted on the question of how to represent "angle" measurements by a selfadjoint operator[10,11]. Additional unnecessary confusion in the case of semibounded number and quantum optical phase was generated by the ignorance or lack of acceptance of generalized ("POVM") observables. On the positive side, an influential paper by Judge[12] produced a relation (for the arc metric), and conjectured an improvement, which was proved shortly afterwards[13]. In this context the role of ground state problems for finding optimal bounds, which is also the basis of our methods, seems to have appeared for the first time[14]. The appearance of the chordal metric grew out of the approach of avoiding the "angle" problem, replacing θ by the two selfadjoint operators $\cos\theta$ and $\sin\theta$.

B.   Measures of uncertainty and error

Both in preparation and in measurement uncertainty we have to assess the difference of probability distributions: For preparation uncertainty it is the difference from a sharp distribution concentrated on a single point. This is also the basis of calibration error assessment. For metric error we also need to express the distance of two general distributions. We want to express the distance between distributions on the same scale as the distance of points. For example, in the case of position and momentum errors, for all error measures one considers ΔQ to be measured in length units and ΔP in momentum units, and this is

satisfied by the choice of standard deviation for these measures.

So let us assume that the outcomes of some observable (represented by a POVM, a positive operator valued measure) lie in a space X with metric d. For real valued quantities like a single component of position or momentum this usually means $X = \mathbb{R}$ and $d(x,y) = |x - y|$.

We now extend the distance function on the points to a distance between a probability measure μ on X and a point $x \in X$. It will just the be the mean distance from x:

$$d_\alpha(\mu, x) = \left(\int \mu(dy) d(y, x)^\alpha \right)^{\frac{1}{\alpha}}. \tag{1}$$

Here the exponent $\alpha \in [1, \infty)$ gives some extra flexibility as to how large deviations are weighted relative to small ones. The αth root ensures that the result is still in the same units as d, and also that for a point measure μ concentrated on a point y we have $d_\alpha(\mu, x) = d(y, x)$ for all α. It is also true for all α that $d_\alpha(\mu, x) = 0$ happens only for the point measure at x. We will later mostly choose $\alpha = 2$ and drop the index α; in this case $d_2(\mu, x)$ is the root mean square distance from x for points distributed according to μ.

With this measure of deviation of a distribution μ from the point we can introduce the generalized standard deviation,

$$d_\alpha(\mu, *) = \min_{x \in X} d_\alpha(\mu, x). \tag{2}$$

Note that for $\alpha = 2$, $X = \mathbb{R}$ and $d(x, y) = |x - y|$ this recovers exactly the usual standard deviation, with the minimum being attained at the mean of μ. The symbol $*$ is just a reminder of the minimization, and emphasizes that $d_\alpha(\mu, *)$ is just the distance of μ from the set of point measures.

For this interpretation to make sense we must also let the second argument of d_α be a general probability distribution ν, resulting in a metric on the set of probability measures. The canonical definition here is the transport distance[15]

$$d_\alpha(\mu, \nu) = \inf_\gamma \left\{ \left(\int \gamma(dx\, dy) d(x, y)^\alpha \right)^{\frac{1}{\alpha}} \ \middle| \ \mu \triangleleft \gamma \triangleright \nu \right\}, \tag{3}$$

where "$\mu \triangleleft \gamma \triangleright \nu$" is a shorthand for γ, the variable in this infimum, being a "coupling" of μ and ν, i.e., it is a measure on $X \times X$ with μ and ν as its marginal distributions. One should think of γ as a plan for converting the distribution μ into ν, maybe for some substance rather than for probability. The cost of transferring a mass unit from x to y is supposed to be $d(x, y)^\alpha$, and the plan γ records just how much mass is to be moved from x to y.

The marginal property means that the initial distribution is μ and the final one ν. Then $d_\alpha(\mu,\nu)^\alpha$ is the optimized cost. When the final distribution is a point measure, there is not much to plan, and we recover (1). Therefore there is little danger of confusion in using the same symbol for the metrics of points and of probability measures.

Now we can use these notions of spread and distance for expressing uncertainties related to observables A, B with outcome spaces X and Y, each with a suitably chosen metric and error exponent. Let us denote by ρ_A the probability measure of outcomes in X upon measuring A on systems prepared according to ρ. To express *preparation uncertainty*, let us consider the set PU of generalized variance pairs

$$\mathsf{PU} = \left\{ \big(d_\alpha(\rho_A,*)^\alpha,\, d_\beta(\rho_B,*)^\beta\big) \mid \rho \text{ a state}\right\}. \tag{4}$$

A preparation uncertainty relation is some inequality saying that the uncertainty region does not extend to the origin: the two deviations cannot both be simultaneously small. If this set is known, we consider it as the most comprehensive expression of preparation uncertainty. Its description by inequalities for products or weighted sums or whatever other expression is a matter of mathematical convenience, and we will, of course, develop appropriate expressions. A lower bound for the product is useful *only* for position and momentum and its mathematical equivalents. In this case the dilatation invariance $(q,p) \mapsto (\lambda q, \lambda^{-1}q)$ forces the uncertainty region to be bounded by an exact hyperbola. But if one of the observables considered can take discrete values, the set will reach an axis, making every state-independent lower bound on the product trivial.

We should note that the set (4) is in general not convex, and can have holes (for examples, see[16]). However, in order to express lower bounds, the essence of uncertainty, it makes no difference if we fill in these holes, and include with every point also those for which both coordinates are larger or the same. The resulting set, the 'monotone hull' PU^+ of PU, is bounded below by the graph of a non-increasing function, the *tradeoff curve* (see Fig. 1 for examples). It still need not be convex in general, but we will see that convexity holds in the examples we study.

For *measurement uncertainty* we again consider two observables (POVMs) A, B with the same outcomes, metrics and error exponents. Now the question is: can A, B be measured jointly? The claim is, usually, that no matter how we try there will be an error in our implementation. So let A', B' be the margins of some joint measurement with outcomes

$X \times Y$. Then A' must exhibit some errors relative to A, i.e., some output distributions $\rho_{A'}$ must be different from ρ_A. We define as the error of A' with respect to A the quantity

$$d_\alpha(A', A) = \sup_\rho d_\alpha(\rho_{A'}, \rho_A). \tag{5}$$

Note that we are using here a worst case quantity with respect to the input state. This is what we should do for a figure of merit for a measuring instrument. If a manufacturer claims that his device A' will produce distributions ε-close to those of A for any input state, he is saying that $d_\alpha(A', A) \leq \varepsilon$. Making such a claim for just a single state is as useless as advertising a clock which measures "the time 12:00" very precisely (but maybe no other). Now we can look at the uncertainty region

$$\mathsf{MU} = \left\{ \left(d_\alpha(A', A)^\alpha, \, d_\beta(B', B)^\beta \right) \mid A' \lhd\!\!\rhd B' \right\}, \tag{6}$$

where the notation $A' \lhd\!\!\rhd B'$ indicates that A' and B' are jointly measurable (that is, they are margins of some POVM that serves as a joint observable). All general remarks made about the preparation uncertainty region PU also hold for MU.

The supremum in (5) is rather demanding experimentally. Good practice for testing the quality of a measuring device is *calibration*, i.e., testing it on states with known properties, and seeing whether the device reproduces these properties. In our case this means testing the device A' on states ρ whose A-distribution is sharply concentrated around some x, and looking at the spread of the A'-distribution around the same x. We define as the calibration error of A' with respect to A the quantity

$$\Delta_\alpha^c(A', A) = \lim_{\varepsilon \to 0} \sup \left\{ d_\alpha(\rho_{A'}, x) \mid d_\alpha(\rho_A, x) \leq \varepsilon \right\}. \tag{7}$$

Here the limit exists because the set (and hence the sup) is decreasing as $\varepsilon \to 0$. This definition only makes sense if there actually are sufficiently many sharp states for A, so we will use this definition only when the reference observable A is projection valued. Since the calibration states in this definition are also contained in the supremum (5), it is clear that $\Delta_\alpha^c(A', A) \leq d_\alpha(A', A)$, so the set CU of calibration error pairs will generally be larger than MU.

III.   SETTING AND OVERVIEW OF RESULTS

We now consider systems with Hilbert space $\mathcal{H} = L^2(\mathbb{T}, d\theta)$, where $\mathbb{T}$ denotes the unit circle, with $d\theta$ the integration over angle. The notation derives from "torus" and is customary

in group theory. We use it here to emphasize the group structure (of multiplying phases or adding angles mod 2π) but also to avoid a fixed coordinatization such as $\mathbb{T} \cong [-\pi, \pi)$, which would misleadingly assign a special role to the cut point $\pm\pi$. We will refer to $\mathbb{T}$ as our "position space". The corresponding "momentum" space is $\mathbb{Z}$, and changing to the momentum representation in $\ell^2(\mathbb{Z})$ is done by the unitary operator of expanding in a Fourier series. With $\{e_n\}_n \in \mathbb{Z}$ denoting the Fourier basis, this means that

$$(\mathcal{F}\psi)(n) = \langle e_n | \psi \rangle = \frac{1}{\sqrt{2\pi}} \int d\theta \, e^{in\theta} \psi(\theta). \tag{8}$$

We have two natural projection valued observables, the angle (=position, phase) observable Θ taking values on $\mathbb{T}$, and the number (=angular momentum observable) N with values in $\mathbb{Z}$. That is, if $f : \mathbb{T} \to \mathbb{C}$ is some function of the angle variable, $f(\Theta)$ denotes the multiplication operator $(f(\Theta)\psi)(\theta) = f(\theta)\psi(\theta)$, and similarly $g(N)$ denotes the multiplication by a function $g(n)$ in the momentum representation. The outcome proability densities of these observables on an input state ρ are denoted by ρ_Θ and ρ_N, respectively. Thus,

$$\int d\theta \, \rho_\Theta(\theta) \, f(\theta) = \text{tr} \, \rho f(\Theta), \tag{9}$$

$$\sum_n \rho_N(n) \, g(n) = \sum_n \langle e_n | \rho | e_n \rangle \, g(n). \tag{10}$$

Then the basic claim of preparation uncertainty is that ρ_Θ and ρ_N cannot be simultaneously sharp, and the basic claim of measurement uncertainty is that there is is no observable with pairs of outcomes (θ, n) for which the marginal distributions found on input state ρ are close to ρ_Θ and ρ_N.

In order to apply the ideas of the previous section, we need to choose a metric in each of these spaces. For discrete values ($X = \mathbb{Z}$) we can naturally take the standard distance or a discrete metric:

$$d_{\text{std}}(n, m) = |n - m| \qquad \text{or}$$

$$d_{\text{dis}}(n, m) = 1 - \delta_{nm}. \tag{11}$$

Similarly, there are two natural choices for angles, depending on whether the basis for the comparison is how far we have to rotate to go from θ to θ' ("arc distance") or else the distance of phase factors $\exp(i\theta)$ and $\exp(i\theta')$ in the plane ("chordal distance"):

$$d_{\text{arc}}(\theta, \theta') = \min_{n \in \mathbb{Z}} |\theta - \theta' - 2\pi n| \qquad \text{or}$$

$$d_{\text{cho}}(\theta, \theta') = \left| e^{i\theta} - e^{i\theta'} \right| = 2 \left| \sin \frac{\theta - \theta'}{2} \right|. \tag{12}$$

The variances based on these bounded metrics will have an upper bound. Since the minimum in (2) makes $d_\alpha(\mu, *)^\alpha$ a concave function of μ, so that we find, by averaging over translates, that the equidistribution has the maximal variance for all translation invariant metrics and all exponents. Both metrics, or rather their quadratic ($\alpha = 2$) variances have been discussed before. The d_{arc}-variance was used by Lévy[17] and Judge[12], the d_{cho}-variance seems to have appeared first in von Mises[18]. In fact, the quadratic chordal variance can also be written as

$$
\begin{aligned}
d_{\mathrm{cho},2}(\mu, *)^2 &= \inf_\alpha \int \mu(d\theta) \left| e^{i\theta} - e^{i\alpha} \right|^2 \\
&= 2\left(1 - \sup_\alpha \mathrm{Re}\ e^{-i\alpha} \int \mu(d\theta) e^{i\theta}\right) \\
&= 2\left(1 - \left|\langle e^{i\theta}\rangle_\mu\right|\right),
\end{aligned}
\tag{13}
$$

which is von Mises' "circular variance". For a review of these choices see Ref. 19. The only property needed in our approach is that the metric should not break the rotation invariance, i.e., it should be a function of the difference of angles. We will therefore use every metric d also as a single variable function, i.e., $d(x) = d(x, 0)$ and $d(x, y) = d(x - y)$. Functions which do not come from a metric have been considered in Ref. 20.

Then we have the following result.

Proposition 1. *For all error exponents and choices of translation invariant metric the three uncertainty regions $PU^+ = MU^+ = CU^+$ coincide. They are depicted for $\alpha = \beta = 2$ in Figure 1. Every point on one of the tradeoff curves belongs to a unique pure state (resp. a unique extremal phase space covariant joint measurement).*

The proof of this Proposition is based entirely on the corresponding proof for standard position and momentum[7]. We will sketch the main steps in the next section, and also show how the computation of the tradeoff curve can be reduced to solving ground state problems for certain Hamiltonians. The detailed features of these diagrams are then developed in the subsequent sections, sorted by the methods employed. The tradeoff curves in Figure 1 for the generalized variances, denoted simply $\Delta^2(N), \Delta^2(\Theta)$, are determined numerically (see Section V). Since the algorithms employed provide optimal bounds, the figures are correct within pixel accuracy (which can be easily pushed to high accuracy). In fact, the problem is very stable, in the sense that near minimal uncertainty implies that the state (or joint observable) is close to the minimizing one. The bounds for this are in terms of the spectral gap of the Hamiltonian and are also discussed in Section V.
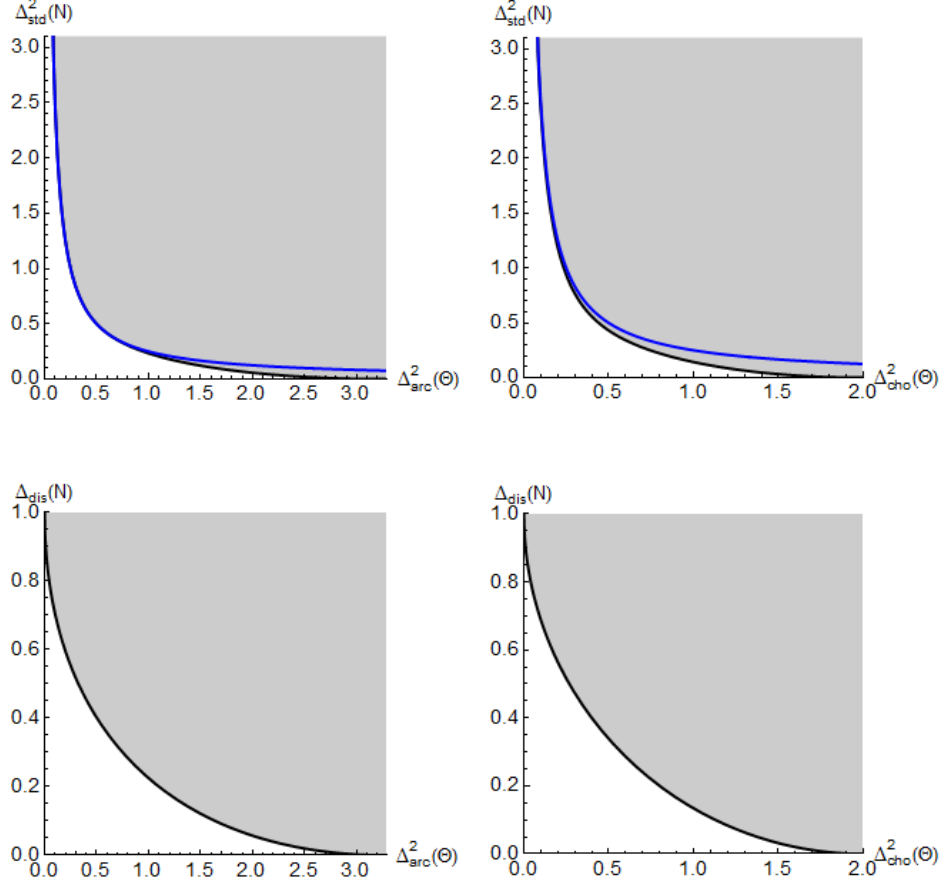
FIG. 1. The uncertainty regions for the four pairs of metrics given by (11) and (12). Each region represents at the same time preparation and measurement uncertainty (for metric or calibration error criterion). For the standard metric on $\mathbb{Z}$ the position-momentum bound is shown for comparison as a blue line. $\Delta_{\text{std}}^2(N)$ is an unbounded quantity. For the others the maximal interval is displayed.

However, the only case in which the exact tradeoff curves can be described in closed form is (see Section VI B)

$$\Delta_{\text{dis}}(N) \geq 1 - \frac{1}{2}\sqrt{\Delta_{\text{cho}}^2(\Theta)\left(4 - \Delta_{\text{cho}}^2(\Theta)\right)} , \tag{14}$$

even though, in all cases, the optimizing states can be expressed explicitly in terms of standard special functions (see Section VI). Therefore simple and explicit lower bounds are of interest. A problem here is that computing the variances for some particular state always produces a point inside the shaded area, i.e., an upper bound to the lower bound represented by the tradeoff curve. This is useless for applications, so in Section VII we

develop a procedure proving lower bounds, and thus correct (if suboptimal) uncertainty relations.

IV. PHASE SPACE SYMMETRY AND REDUCTION TO A GROUND STATE PROBLEM

In this section we briefly sketch the arguments leading to the equality of preparation und measurement uncertainty regions. The full proof is directly parallel to the one given in Ref. 7 for position and momentum. The basic reference for phase space quantum mechanics is Ref. 21. The theory there is developed for phase spaces of the form $\mathbb{R}^n \times \mathbb{R}^n$, but all results we need here immediately carry over to the general case $X \times \widehat{X}$, where X is a locally compact abelian group and $\widehat{X}$ is its dual, in our case $X = \mathbb{T}$ and $\widehat{X} = \mathbb{Z}$. A systematic extension of Ref. 21 to the general case, including the finer points, is in preparation in collaboration with Jussi Schultz.

A. Covariant phase space observables

The phase space in our setting is the group $\Omega = \mathbb{T} \times \mathbb{Z}$. We join the position translations and the momentum translations to phase space translations, which are represented by the displacement or *Weyl operators*

$$(W(n,\theta)\psi)(x) = e^{-\frac{1}{2}n\theta} e^{-inx} \psi(x - \theta). \tag{15}$$

These operators commute up to a phase, so that the operators $T_\omega(A) = W(\omega)^* A W(\Omega)$, i.e., the action of the Weyl operators on bounded operators $A \in \mathcal{B}(\mathcal{H})$, is a representation of the abelian group Ω. A crucial property is the square integrability of the matrix elements of the Weyl operators, which we will use in the form that for any two trace class operators ρ, σ on $\mathcal{H}$ the formula

$$\int d\omega \operatorname{tr}\big(\rho\, T_\omega(\sigma)\big) = \operatorname{tr} \rho \, \operatorname{tr} \sigma, \tag{16}$$

where $\int d\omega = \sum_n \int \frac{d\theta}{2\pi}$. Hence, when both ρ and σ are density operators the integrand in this equation is a probability density on Ω. In summary[21]:

Proposition 2. *Every density operator σ serves an observable F_σ with outcomes ω in the*

phase space Ω, via

$$F_\sigma(A) = \int_{\omega \in A} d\omega \; T_\omega(\sigma). \tag{17}$$

Here σ figures as the operator valued Radon-Nikodym density with respect to $d\omega$ of F_σ at the origin, and by translation also at arbitrary points. The observables obtained in this way are precisely the covariant ones, i.e., those satisfying the equation $T_\omega(F(A)) = F(A - \omega)$.

For the discussion of uncertainties we need the margins of such observables. One can guess their form from a fruitful classical analogy[21], by which the integrand in (16) can be read as a convolution of ρ and σ. For classical probability densities on a cartesian product it is easily checked that the margin of the convolution is the convolution of the margins. The same is true for operators, only that the margins of a density operator are the classical distributions ρ_Θ and ρ_N described above. If p is a probability density on phase space we will also denote by p_Θ and p_N the respective margins on $\mathbb{T}$ and $\mathbb{Z}$. In particular, for the output distribution of the covariant observable F_σ, i.e., $p(\theta, n) = \operatorname{tr} \rho W(\theta, n)^* \sigma W(\theta, n)$, One checks readily checks the marginal relations

$$p_N = \rho_N * \sigma_N, \tag{18}$$

$$p_\Theta = \rho_\Theta * \sigma_\Theta, \tag{19}$$

where "$*$" means the convolution of probability densities on $\mathbb{Z}$ and $\mathbb{T}$. Since this is the operation associated with the sum of independent random variables we arrive at the following principle:

Corollary 1. *Both margins of a covariant phase space measurement F_σ can be simulated by first making the corresponding ideal measurement on the input state, and then adding some independent noise, which is also independent of the input state. The distribution of this noise is the corresponding margin of the density operator σ.*

This principle is responsible for the remarkable equality of the preparation uncertainty region PU^+ and the measurement uncertainty regions MU^+ and CU^+. Indeed the added noise is what distinguishes the margins of an attempted joint measurement from an ideal measurement, and this has precisely the distribution relevant for preparation uncertainty for σ.

B. MU and CU: Reduction to the covariant case

While this principle explains quite well what happens in the case of covariant observables F_σ, Proposition 1 makes no covariance assumption. The key for reducing the general case is the observation that our quality criteria in terms of $d(\Theta', \Theta)$ do not single out a point in phase space. Thus, let $\mathcal{M}(\varepsilon_1, \varepsilon_2)$ be the set of observables whose angle margin $F_\mathbb{T}$ is ε_1-close to the ideal observable, $d(F_\mathbb{T}, \Theta) \leq \varepsilon_1$, and whose number margin satisfies $d(F_\mathbb{Z}, N) \leq \varepsilon_2$; then this set is closed under phase space shifts, i.e., is unchanged when we replace F by F' where

$$F'(A) = T_\omega(F(A + \omega)) \tag{20}$$

Note that the fixed points of all these transformations are precisely the covariant observables. The second point to note is that the set $\mathcal{M}(\varepsilon_1, \varepsilon_2)$ is convex, because the worst case error of an average in a convex combination is smaller than the average of the worst case errors. It is also a compact set in a suitable weak topology. This is in contrast to the set of all observables: Since there are arbitrarily large shifts in $\mathbb{Z}$, we can shift an observable to infinity such that the probabilities for all fixed finite regions go to zero. The weak limit of such observables would be zero or, in an alternative formulation, would acquire some weight on points at infinity (a compactification of $\mathbb{Z}$). For such a sequence, however, the errors would also diverge. It is shown in Ref. 7 that this suffices to ensure the compactness of $\mathcal{M}(\varepsilon_1, \varepsilon_2)$. Then the Markov-Kakutani Fixed Point Theorem (Ref. 22, Thm. V.10.6) ensures that $\mathcal{M}(\varepsilon_1, \varepsilon_2)$ contains a common fixed point of all the transformations (20).

In summary:

Proposition 3. *For every joint observable F on the phase space Ω there is a covariant one for which the errors are at least as good.*

Therefore for determining MU^+ we can just assume the observable in question to be covariant, implying the very simple form of the margins described above. The argument for CU^+ is the same.

C. The post-processing Lemma: $\mathsf{CU}^+ = \mathsf{MU}^+$

We have noted that, in general, $\Delta_\alpha^c(A', A) \leq d_\alpha(A', A)$, because for calibration the worst case analysis is done over a much smaller set of states. Indeed, it is easy to construct examples

of observable pairs where the inequality is strict. There is a general result, Ref. 16, Lemma 8, however, which implies equality. The condition is that A' arises from A by classical, possibly stochastic post-processing. That is, we can simulate A' by first measuring A, and then adding noise or, in other words, generating a random output by a process which may depend on the measured A-value. The noise is described then by a transition probability kernel $P(x, dy)$ for turning the measured value x into somewhere in the set $dy \subset X$. P is thus a description of the noise, and its relevant size is given by the formula

$$\Delta_\alpha^c(A', A) = d_\alpha(A', A) = \left(\underset{A}{\text{ess-sup}} \underset{x \in X}{\int} P(x, dy) \, d(x, y)^\alpha \right)^{1/\alpha}. \tag{21}$$

Here the A-essential supremum of a measurable function is the supremum of all λ such that the level set $\{x | f(x) \geq \lambda\}$ has non-zero measure with respect to A. This is needed to ensure that $P(x, dy)$ enters this formula only for values x that can actually occur as outputs of A.

D.   The covariant case: $\mathbf{CU}^+ = \mathbf{MU}^+ = \mathbf{PU}^+$

In the case at hand, the noise is independent of x, i.e., P is translation invariant and so is the metric. Therefore, the integral in (21) is simply independent of x. Moreover, we know the distribution of the noise on each margin to be the respective margin of σ, so that the integral is just the α-power deviation of the margin from zero. So we get, for any choice of exponents and translation invariant metrics on $\mathbb{T}$ and $\mathbb{Z}$:

$$\Delta_\alpha^c(\Theta', \Theta) = d_\alpha(\Theta', \Theta) = d_\alpha(\sigma_\Theta, 0), \tag{22}$$

and similarly for N. Note that the last term here is *not* the variance $d_\alpha(\sigma_\Theta, *)$, because the minimization over x in (2) is missing. Indeed, if σ_Θ just had zero variance, i.e, it were a point measure at some point $x \neq 0$, we would get a constant shift of size $d(x)$ between the distributions $\rho_{\Theta'}$ and ρ_Θ, and this would be the errors on the left hand side. So for a fixed σ we can only say that the terms in (22) are $\geq d_\alpha(\sigma_\Theta, *)$. On the other hand, we are looking for *optimal* σ and these will be obtained by shifting σ in such a way that (22) is minimized. Hence, as far as uncertainty diagrams are concerned, we can replace the last term by the α-deviation. This concludes the proof that the three uncertainty diagrams coincide.

E. Minimizing the variances

We now describe the general method to find the tradeoff curve. The idea is to fix some negative slope $-t$ in the diagram, and ask for the lowest straight line with that slope intersecting PU^+. That is, we look at the optimal lower bound $c()$ such that

$$y + tx \geq c(t) \qquad \text{for all } (x, y) \in \mathsf{PU}. \tag{23}$$

Now both coordinates $x = d_\alpha(\sigma_\Theta, 0)^\alpha$ and $y = d_\beta(\sigma_N, 0)^\beta$ are linear functions of σ, so that the left hand side of (23) is just the σ-expectation of some operator, namely

$$y + tx = \operatorname{tr}(\sigma H(t))$$
$$H(t) = d_\mathbb{Z}(N)^\beta + t d_\mathbb{T}(\Theta)^\alpha. \tag{24}$$

Here $d_\mathbb{T}$ and $d_\mathbb{Z}$ are the metrics chosen for these spaces, and we used the notation of writing $f(\Theta)$ for the multiplication operator by $f(\theta)$, and its Fourier transformed counterpart, and also the convention that $d(x) = d(x, 0)$ for a translation invariant metric. The optimal constant $c(t)$ is thus the lowest expectation $\inf H = \inf \operatorname{tr}(\sigma H)$, i.e., its *ground state energy.* Note that for standard position and momentum phase space and $\alpha = \beta = 2$ we get here $H = P^2 + tQ^2$, a harmonic oscillator, and the well-known connection between its ground state and minimum uncertainty.

We will look into these ground state problems later and for now note some general features.

(i) The variable t is positive because we are looking for lower bounds on x and y only. This corresponds in part to taking the monotone closure, and is the reason why we replace PU^+ by PU in (23).

(ii) The best bound on PU obtained in this way is achieved by optimizing over t, i.e.,

$$y \geq \sup_t \{-tx + c(t)\},$$

the Legendre transform of c. This is automatically convex. In other words, the method does not describe PU^+ in general, but its convex hull (the intersection of all half spaces with positive normal containing PU).

(iii) There may be points on the tradeoff curve for the convex hull which do not really correspond to a realizable pair of uncertainties. However, if we take the collection of (t-dependent) ground states, and their variance pairs trace out a continuous curve, we know that the tradeoff curves are the same and the set PU^+ is actually convex and fully characterized by the ground state method.
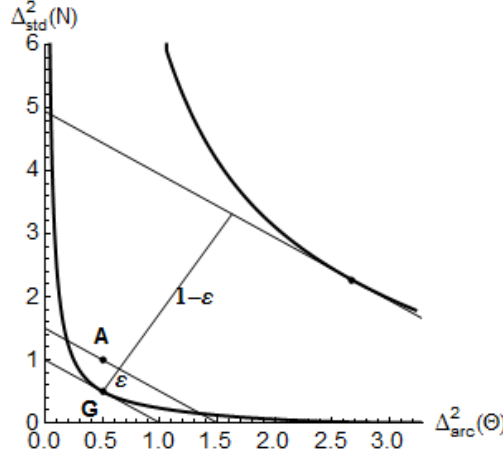

FIG. 2. Applying the estimate (25) in an uncertainty diagram. The lower curve is the tradeoff curve obtained from the ground state problem. The upper curve is formed by the uncertainty pairs of the respective first excited states. Suppose the uncertainty pair A has been found. We want to show that the corresponding state must be close to the minimal uncertainty state corresponding the point G. Drawing the tangent to the tradeoff curve at G and the parallel tangent to the upper curve we find from the diagram that the fidelity of the given state to G must be at least $1 - \varepsilon$.

When the ground state problem for $H(t)$ has a gap, it is known that any state with expectations close to the ground state energy must actually be close to the ground state. More precisely, suppose that H has a unique ground state vector, $H\psi = E_0\psi$, and that the next largest eigenvalue is $E_1 > E_0$. Then $H \geq E_1\mathbb{1} - (E_1 - E_0)|\psi\rangle\langle\psi|$. Now let $E_0 < \langle\phi, H\phi\rangle = E_\phi < E_1$ for some unit vector ϕ. Then by taking the ϕ-expectation of the operator inequality, we get

$$|\langle\phi, \psi\rangle|^2 \geq \frac{E_1 - E_\phi}{E_1 - E_0}. \tag{25}$$

In particular, when $E_\phi \approx E_0$, ϕ must be close in norm to ψ. We can directly apply this principle to the above ground state problems. The basic geometry is described in Fig. 2. This shows that the curve of minimizers is continuous. It will also be useful in showing

explicitly that the minimizers for different choices of metrics are sometimes quite close to each other, or that some simple ansatz for the minimizer is quantitatively good.

V. NUMERICS IN TRUNCATED FOURIER BASIS

Here we only consider the case $d_{\mathbb{Z}} = d_{\mathrm{std}}$ because for the discrete metric on $\mathbb{Z}$ the ground state problem has an elementary explicit solution (see Section VI B). The numerical treatment is easiest in the Fourier basis, or rather in the even eigenspace of the number operator N, for $|N| \leq N_{\max}$. The matrix elements of the relevant Hamiltonians for basis vectors in this range can be written down as simple explicit expressions. From these the numerical version of the Hamiltonian is determined as floating point matrix of the desired precision, for which the ground state and first excited state are determined by standard algorithms. All these steps were carried out in Mathematica. The criterion for the choice of $N_{\max}$ was that the highest-n components of the eigenvectors found should be negligible at the target accuracy. The target accuracy was mostly 5 digits with computations done in machine precision with $N_{\max} = 80$, but was chosen larger for getting a reliable estimate of the separation of the different state families.

All computations must be considered elementary and highly efficient, even at high accuracy. None of the diagrams in this paper takes computation time longer than a keystroke. It is therefore hardly of numerical advantage to implement the analytic solutions of Section VI, not in computation time and even less in programming and verification time.

Perhaps the only surprise in this problem is that for the two different metrics d_{arc} and d_{cho} the minimizing state families are so close. Since $V_{\mathrm{cho}} \leq V_{\mathrm{arc}}$, the ground state problems for H_{cho} and H_{arc} are related. Perturbatively one sees that the ground state energies are indeed similar, up to the expectation of $t(V_{\mathrm{arc}}(\theta) - V_{\mathrm{cho}}(\theta))$. The stability statement at the end of Section IV E then implies that the corresponding ground states are also similar. However, direct comparison gives a norm bound, which is rather better than these arguments indicate:

$$||\psi_{\mathrm{arc}}(t) - \psi_{\mathrm{cho}}(t)|| \leq 0.145 \tag{26}$$

for all t, corresponding to a fidelity $\geq .98$. This still does not quite reflect the similarity of these two state families: When we allow the t-arguments to differ, we get a much better approximation. To make this precise consider the orbits $\Omega_{\mathrm{arc}} = \{e^{i\alpha}\psi_{\mathrm{arc}}(t)|t > 0, \alpha \in \mathbb{R}\}$,

and an analogously defined Ω_{cho}. For sets in Hilbert space we use the Hausdorff metric, so that $d_H(\Omega_1, \Omega_2) < \varepsilon$ means that for every point in one set there is an ε-close one in the other. Then one easily gets

$$d_H\big(\Omega_{\text{arc}}, \Omega_{\text{cho}}\big) \leq .028 \tag{27}$$

Consequently, there is really only one diagram representing the family of minimal uncertainty states, which we show in Fig. 3.
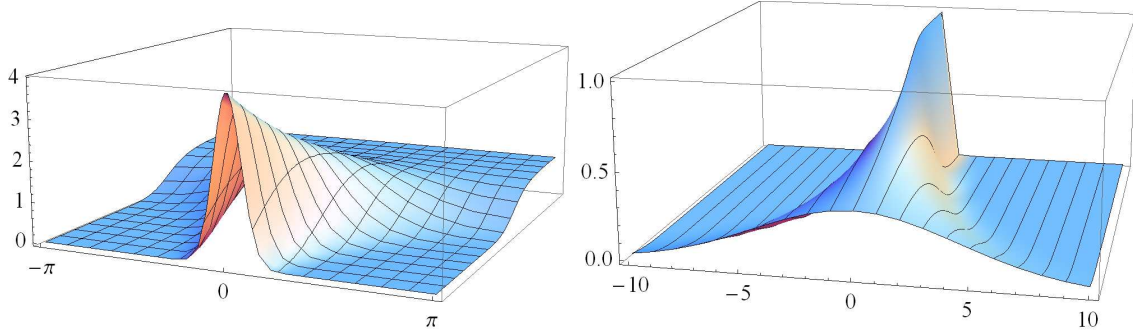

FIG. 3. Minimum uncertainty wave functions (left) and Fourier coefficients (right) for the case $(d_{\text{cho}}, d_{\text{std}})$ as a function of the family parameter t from (24), plotted as the depth dimension. Note that the surface in the diagram on the right is only an aid for better 3D visualization — only the embedded lines have real significance.

VI. EXACT GROUND STATES

A. Schrödinger operator case

With $d_{\mathbb{Z}} = d_{\text{std}}$ and $\beta = 2$, the ground state problem becomes an instance of the Schrödinger operator eigenvalue problem. In fact, writing $V = d_{\mathbb{Z}}^\alpha(\Theta)$, the optimal constant $c(t)$ for a given t is the smallest value of λ such that the differential equation

$$-\psi''(\theta) + tV(\theta)\psi(\theta) = \lambda\psi(\theta) \tag{28}$$

has a solution on $[-\pi, \pi]$ satisfying the boundary conditions $\psi(-\pi) = \psi(\pi) = \psi'(-\pi) = \psi'(\pi) = 0$. By the general theory, we know that the (unique) solution $\psi = \psi_\alpha$ has no zeros, can be chosen to depend smoothly on t (by perturbation theory), and can be chosen to be positive and even (by parity invariance).

18

Hence, we are in fact looking for an even solution ψ of (28) with $\psi'(\pi) = 0$. At $t = 0$ this vector is just $\psi_\alpha(\theta) = 1/\sqrt{2\pi}$, i.e. a constant. For the two choices d_{arc} and d_{cho}, the solutions ψ_α are known special functions; we now proceed to describe them in some detail.

1. $V_{\mathbb{Z}}(\theta) = d_{\mathrm{arc}}(\theta)^2 = \theta^2$

The general even solution of (28) is given by a hypergeometric function[23]:

$$\psi(\theta) = N^{-1} e^{-\frac{1}{2}\sqrt{t}\,\theta^2} \,_1\mathrm{F}_1\left(\frac{1}{4}\left(1 - \frac{\lambda}{\sqrt{t}}\right); \frac{1}{2}; \sqrt{t}\theta^2\right), \tag{29}$$

where N is the normalisation factor. The boundary condition $\psi'(\pi) = 0$ now picks out the eigenvalues λ for every t (see Fig. 4), of which the smallest is the desired constant $c(t)$. The condition can be expressed by using the standard differentiation formulas for the hypergeometric functions:

$$\left(1 - \frac{\lambda}{\sqrt{t}}\right) \,_1\mathrm{F}_1\left(1 + \frac{1}{4}\left(1 - \frac{\lambda}{\sqrt{t}}\right); \frac{3}{2}; \pi^2\sqrt{t}\right) = \,_1\mathrm{F}_1\left(\frac{1}{4}\left(1 - \frac{\lambda}{\sqrt{t}}\right); \frac{1}{2}; \pi^2\sqrt{t}\right)$$

However, as far as we could see, the theory of hypergeometric functions seems to offer little help for solving it, or for evaluating the normalization constant N or the Fourier coefficients. Perhaps an elementary expression for $c(t)$ is too much to hope for, since already in much simpler problems, e.g., a particle in a box, where the pertinent transcendental equations involve only trigonometric and linear functions, no "explicit" solution can be given either.
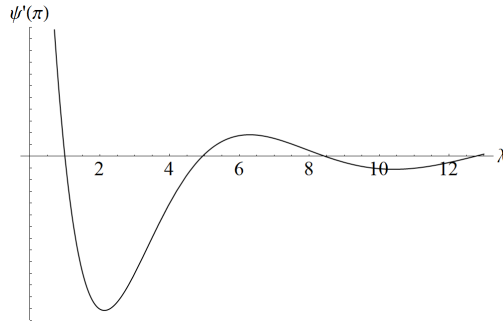

FIG. 4. The derivative $\psi'(\pi)$ after (29), for $t = 1$, as a function of the eigenvalue parameter λ. The zeros of this function determine the eigenvalues, of which the lowest gives the constant $c(1)$.

2. $V(\theta) = d_{\mathrm{cho}}(\theta)^2 = 2(1 - \cos\theta)$

In the case of d_{cho} (distance through the circle), the equation (28) is just the Mathieu equation up to scaling $\theta \mapsto \theta/2$. The even periodic solutions correspond to $-4(2t - \lambda) =$ $a_r(-4t)$, where the $a_r(q)$, $r = 0, 1, \ldots$ are called Mathieu characteristic values[23]. Our ground state eigenvalues are therefore

$$c(t) = 2t + a_t/4,$$

where we have used the shorthand $a_t = a_0(-4t)$. Since a_0 is implemented in e.g. Mathematica, we can easily determine the values numerically. The corresponding solutions are given by

$$\psi_t(\theta) = \frac{1}{N}\mathrm{ce}_0\left(a_t; -4t; \frac{\theta}{2}\right)$$

where ce_0 denotes the lowest order first kind solution of the ordinary Mathieu equation, and N is again the normalisation factor. We note that the Fourier coefficients $\widehat{\psi}_t$ are explicit functions of a_t and t, determined by the recurrence relations. Up to second order, we have

$$\psi_t(\theta) = \frac{1}{N}\left(1 - \frac{a_t}{4t}\cos\theta + \left[\frac{(a_t - 4)a_t}{16t^2} + 8t\right]\cos 2\theta + O(\cos 3\theta)\right)$$

The relevance of the Mathieu functions in the context of circular uncertainty relations has been noted e.g. in Ref. 24.

B. Discrete metric case

With $d_{\mathbb{Z}} = d_{\mathrm{dis}}$, we have the eigenvalue equation

$$\left(\mathbb{I} - |\phi_0\rangle\langle\phi_0| + tV\right)\psi = \lambda\psi, \tag{30}$$

with V as in the previous section, and $\phi_0(\theta) = 1/\sqrt{2\pi}$ the constant function. This allows us to solve for ψ:

$$\psi_t(\theta) = \frac{A}{1 - \lambda + tV(\theta)}, \tag{31}$$

where $A > 0$ is the normalization constant. Inserting this into (30) gives the consistency condition

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{d\theta}{1 - \lambda + tV(\theta)} = 1. \tag{32}$$

The smallest positive solution $\lambda = c(t)$ of this equation will give us the desired bound.

However, we can also proceed more directly by using just the functional form (31), which we can further simplify to the one-parameter family $\psi(\theta) = A/(\mu + V(\theta))$, with a single parameter μ. We then have to solve three integrals:

$$I_1(\mu) = \int_{-\pi}^{\pi} \frac{d\theta}{\mu + V(\theta)}$$

$$I_2(\mu) = \int_{-\pi}^{\pi} \frac{d\theta}{(\mu + V(\theta))^2} = -\frac{dI_1(\mu)}{d\mu}$$

$$I_3(\mu) = \int_{-\pi}^{\pi} \frac{d\theta \, V(\theta)}{(\mu + V(\theta))^2} = I_1(\mu) - \mu I_2(\mu).$$

Then the pair of variances

$$\Delta_{\mathrm{dis}}(N) = 1 - \frac{I_1(\mu)^2}{2\pi I_2(\mu)} \tag{33}$$

$$\Delta^2(\Theta) = \frac{I_3(\mu)}{I_2(\mu)} \tag{34}$$

lies on the tradeoff curve. One can check that this is consistent with the Legendre transform picture, i.e., condition (32) in the form

$$I_1\left(\frac{1 - c(t)}{t}\right) = 2\pi t \tag{35}$$

and its derivative and the parameter identification $\mu = (1 - c)/t$.

Now for the arc metric we have $V(\theta) = \theta^2$ and

$$I_1(\mu) = \frac{2}{\sqrt{\mu}} \arctan\left(\frac{\pi}{\sqrt{\mu}}\right). \tag{36}$$

Trying to eliminate μ from (33) leads to a transcendental equation, so one cannot give a closed inequality involving just the variances.

For the chord metric we have $V(\theta) = 2(1 - \cos\theta)$ and

$$I_1(\mu) = \frac{2\pi}{\sqrt{\mu(\mu + 4)}}. \tag{37}$$

In this case one can easily eliminate μ from (33), and the tradeoff is explicitly described by equation (14).

VII.   ANALYTIC LOWER BOUNDS

In this section we establish a variational method for proving uncertainty relations by applying such bounds for the ground state problem. Of course, variational methods for the

21

ground state problem are well-known. Basically they amount to choosing some good trial state, and evaluating the energy expectation: This will be an upper bound on the ground state energy, and it will be a good one if we have guessed well. However, it is notoriously difficult to find lower bounds on the ground state energy. The idea for finding such bounds is via the following Lemma:

Lemma 1. *Let V be a 2π-periodic real valued potential, and H the Schrödinger operator $H\psi = -\psi'' + V\psi$ with ground state energy $E_0(H)$. Consider a twice differentiable periodic function ϕ, which is everywhere > 0. Then the ground state energy of H is larger or equal to*

$$E_\phi = \min_\theta \left\{ V(\theta) - \frac{\phi''(\theta)}{\phi(\theta)} \right\}. \tag{38}$$

Proof. Let

$$\widetilde{V}(x) = \phi''(x)/\phi(x) + E_\phi \tag{39}$$

and $\widetilde{H}$ the Schrödinger operator with this potential. Then ϕ is an eigenfunction of $\widetilde{H}$ with eigenvalue E_ϕ, and since it was assumed to be positive, it has no nodes and must hence be the ground state eigenfunction. On the other hand, because $E_\phi \leq V(x) - (\widetilde{V}(x) - E_\phi)$, we have $\widetilde{V} \leq V$ and hence $\widetilde{H} \leq H$. By the Rayleigh-Ritz variational principle[25] this implies the ordering of the ground state energies, i.e., $\widetilde{E} = E_0(\widetilde{H}) \leq E_0(H)$. □

Finding a ϕ which gives a good bound is usually more demanding than finding a good L^2-approximant for the ground state, because of the highly discontinuous expression ϕ''/ϕ and the infimum being taken over the whole interval. In particular, the approximate eigenvectors obtained by other methods may perform poorly, even give negative lower bounds on a manifestly positive operator.

The positivity of ϕ may be ensured by setting $\phi(\theta) = \exp f(\theta)$; then one has to minimize $V - (f'' + (f')^2))$.

We now consider the combinations of metrics $(d_{\mathrm{std}}, d_{\mathrm{arc}})$ and $(d_{\mathrm{std}}, d_{\mathrm{cho}})$, also comparing the results with existing uncertainty relations found in the literature. This demonstrates how our systematic approach relates to many existing (seemingly ad hoc) uncertainty relations.

One remark should be made concerning the comparison with the literature: The uncertainty measure used for the number operator is practically always taken to be the usual standard deviation, which can be different from $\Delta_{\mathrm{std}}(N)$ since in the latter case the infimum

is taken only over the set of integers. In general we have

$$\Delta_{\text{std}}(N)^2 \geq \sum_{n \in \mathbb{Z}} (n - \text{tr}(\rho N))^2 \rho_N(n),$$

where the right-hand side is the usual standard deviation, which takes a distribution on the integers as a distribution on the reals, which is supported by the integers. Due to the above inequality any uncertainty tradeoff involving the usual standard deviation also implies the same relation for $\Delta_{\text{std}}(N)$.

A. Case $(d_{\text{std}}, d_{\text{arc}})$

The literature on this case begins with the observation that the standard uncertainty relation does not hold and needs to be modified; Judge[12,26] showed in 1963 that the following tradeoff relation

$$\Delta_{\text{std}}(N)\Delta_{\text{arc}}(\Theta) \geq c \left(1 - (3/\pi^2)\Delta_{\text{arc}}(\Theta)^2\right) \tag{40}$$

holds with $c = 0.16$, and conjectured the same with $c = 1/2$. The conjecture was quickly proved in Ref. 13 using the Lagrange multiplier method where $\Delta_{\text{dis}}(N)$ is minimised under the constraint of fixed $\Delta_{\text{arc}}(\Theta)^2$, and in Ref. 14 and 27 by showing that the admissible pairs $(\Delta_{\text{std}}(N)^2, \Delta_{\text{arc}}(\Theta)^2)$ lie above the tangent lines of the curve determined by the equality in (40). Both methods are essentially equivalent to our approach, and explicitly involve the same eigenvalue problem. In Ref. 27 the bound leading to (40) with the optimal constant was obtained using special properties of the confluent hypergeometric function.

We first show how the above Lemma can easily be applied to derive (40) with the optimal constant $c = 1/2$. An essentially identical procedure works also in other cases below. The potential is $V(\theta) = \theta^2$, and we have label the uncertainty pairs as $(x, y) = (\Delta_{\text{arc}}(\Theta)^2, \Delta_{\text{std}}(N)^2)$. As the simplest ansatz we take $f(\theta) = \log \phi(\theta)$ even, hence a polynomial in θ^2, which take as quadratic. The boundary condition $f'(\pi) = 0$ then leaves the one-parameter family

$$f(\theta) = -\frac{a}{2}\theta^2\left(1 - \frac{\theta^2}{2\pi^2}\right), \tag{41}$$

where $a \in \mathbb{R}$ is to be optimised later. The bound given by the Lemma on the uncertainty pair (x, y) is then

$$y + tx \geq E_0 \geq E_\phi = \inf_\theta\left\{t\theta^2 - f''(\theta) - f'(\theta)^2\right\}$$

$$= a + \inf_\theta\left\{\left(t - a^2 - \frac{3a}{\pi^2}\right)\theta^2 + \frac{2a^2}{\pi^2}\theta^4 - \frac{a^2}{\pi^4}\theta^6\right\}.$$

23

This inequality is valid for any $a > 0$ and $t > 0$. We choose t so that the linear term in ξ vanishes, i.e., $t = a^2 + 3a\pi^{-2}$. The remaining polynomial is then positive because of $\theta \leq \pi$, and hence takes its minimum at $\xi = 0$. Therefore,

$$y + \left(a^2 + \frac{3a}{\pi^2}\right) x \geq a. \tag{42}$$

The optimal value here is $a = (\pi^2 - 3x)/(2\pi^2 x)$. Note that this is alwas positive, because the equidistribution has the largest variance, namely $x = \pi^2/3$. Substituting the optimal a in (42) we get exactly (40).

B.   Case $(d_{\mathrm{std}}, d_{\mathrm{cho}})$

We first recall from (13) that $|\langle e^{i\theta}\rangle_{\rho_\Theta}| = 1 - \Delta_{\mathrm{cho}}(\Theta)^2/2$. Hence, the von Mises "circular variance" is associated with the sine and cosine operators $\sin\Theta$ and $\cos\Theta$, introduced by Carruthers, Nieto, Louisell, Susskind, Glogower, and others to study the "quantum phase problem"[19,28,29]. The idea was to replace the singular commutator $[N, \Theta]$ by the well-defined relations

$$[N, \sin\Theta] = i\cos\Theta, \qquad\qquad [N, \cos\Theta] = -i\sin\Theta. \tag{43}$$

Combining the usual Robertson type inequalities associated with these commutators, they obtained (Ref. 29, eq. (4.11)) the tradeoff relation

$$\Delta_{\mathrm{std}}(N)\Delta_{\mathrm{cho}}(\Theta) \geq \frac{1 - \frac{1}{2}\Delta_{\mathrm{cho}}(\Theta)^2}{2\left[1 - \frac{1}{4}\Delta_{\mathrm{cho}}(\Theta)^2\right]^{1/2}}, \tag{44}$$

expressed here in quantities relevant for our discussion. It was shown by Jackiw[30] that there are no states for which this inequality is saturated, i.e., this bound is not sharp. It is interesting to note that by replacing the square root term with its trivial upper bound 1, we get

$$\Delta_{\mathrm{std}}(N)\Delta_{\mathrm{cho}}(\Theta) \geq \frac{1 - \frac{1}{2}\Delta_{\mathrm{cho}}(\Theta)^2}{2}, \tag{45}$$

which is just the version of Judge's bound (40) for this metric. Other lower bounds were studied relatively recently[24] by using approximations of the Mathieu functions associated with the exact tradeoff curve.

We first show how (45) can be obtained using Lemma 1 by applying the same procedure as above. Interestingly, the relevant trial states are exactly the ones saturating the Robertson

inequality for the first commutator in (43), that is, we take $f(\theta) = a\cos\theta$. Then the resulting variational expression is a function of the variable $\cos\theta$, and hence of the potential V:

$$y + tx \geq \inf_V\left\{a - \left(t - \frac{a}{2} - a^2\right)V + \frac{a^2}{4}V^2\right\} \tag{46}$$

Again it is a good choice to take t so that the first order term in V vanishes, so that the remaining infimum is attained at $V = 0$. This gives $t = a/2 + a^2$ and

$$y \geq a - \left(\frac{a}{2} + a^2\right)x = \frac{(2-x)^2}{16x}. \tag{47}$$

where at the last equality we have substituted the optimal value $a = (2-x)/(4x)$. On taking the square root this is (45).

In order to obtain analytic bounds better than the Carruthers-Nieto tradeoff (44), we apply our method with a trial function which is second order in $\cos\theta$: We take $f(\theta) = a\cos\theta + b(\cos\theta)^2$. The expression to be minimized over θ can still be written as a polynomial in the potential, and numerical inspection suggests once again that it is a good idea to choose the parameters t and b so that coefficients of V and V^2 vanish. This gives linear equations for t and b, and the resulting polynomial has its unique minimum, namely a, at $V = 0$. The analogue of (47) is then

$$y + \frac{a\left(8a^2 + 5a + 2\right)}{8a + 4}x \geq a. \tag{48}$$

Optimizing a now leads to a third order algebraic equation for which the Cardano solution gives a useless expression in terms of roots. If one just wants the tradeoff curve, the solution is actually not necessary. Defining the coefficient of x as a function $g(a)$, so that $y + g(a)x \geq y$. Optimality requires $xg'(a) = 1$, so we get the tradeoff curve in parametrized form $a \mapsto (1/g'(a), a - g(a)/g'(a))$.

VIII.   OUTLOOK

The methods employed in this paper for obtaining preparation uncertainty bounds can be applied to a large variety of similar problems. However, the derivation of measurement uncertainty bounds relied entirely on the theorem that phase space symmetry makes the two coincide. It is therefore no surprise that the case of positive number and phase seems much harder to tackle, and the exact uncertainty region is yet to be determined although (non-strict) uncertainty bounds have recently been proven[31]. Independent efficient methods

for obtaining sharp bounds for measurement uncertainty so far have not been found, and it would be highly desirable to find such methods. A possible substitute might be a proof of the conjecture that measurement uncertainty is always larger than preparation uncertainty. Although this inequality must be strict in general, in that way the easily computed preparation uncertainty bounds would automatically be valid (but usually suboptimal) measurement uncertainty bounds. However, the only evidence for supporting such a conjecture is the comparison of cases where either kind of uncertainty vanishes, so such a result is perhaps too much to hope for.

ACKNOWLEDGEMENTS

REFERENCES

[1] E. Kennard, "Zur Quantenmechanik einfacher Bewegungstypen," Zeitschr. Phys. **44**, 326–352 (1927).

[2] H. Robertson, "The uncertainty principle," Phys. Rev. **34**, 163–164 (1929).

[3] W. Heisenberg, "Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik," Zeitschr. Phys. **43**, 172–198 (1927).

[4] R. Werner, "The uncertainty relation for joint measurement of position and momentum," Quant. Inform. Comput. **4**, 546–562 (2004), quant-ph/0405184.

[5] P. Busch and D. Pearson, "Universal joint-measurement uncertainty relation for error bars," Journal of Mathematical Physics **48**, 082103 (2007).

[6] P. Busch, P. Lahti, and R. Werner, "Proof of Heisenberg's error-disturbance relation," Phys. Rev. Lett. **111**, 160405 (2013).

[7] P. Busch, P. Lahti, and R. Werner, "Measurement uncertainty relations," J. Math. Phys. **55**, 042111 (2014).

[8] J. Biniok, P. Busch, and J. Kiukas, "Uncertainty in the context of multislit interferometry," Phys. Rev. A **90**, 022115 (2014).

[9] S. Tanimura, "Complementarity and the nature of uncertainty relations in Einstein–Bohr recoiling slit experiment," Quanta **4**, 1 (2015).

[10] D. Judge and J. Lewis, "On the commutator $[L_z, \varphi]$," Phys. Lett. **5**, 190 (1963).

[11] K. Kraus, "Remark on the uncertainty between angle and angular momentum," Zeitschrift für Physik **188**, 374 (1965).

[12] D. Judge, "On the uncertainty relation for L_z and φ," Phys. Lett. **5**, 189 (1963).

[13] A. Evett and H. Mahmoud, "Uncertainty relation for angle variables," Nuovo Cimento **38**, 295 (1965).

[14] L. Schotsmans and P. V. Leuven, "Numerical evaluation of the uncertainty relation for angular variables," Nuovo Cimento **39**, 776 (1965).

[15] C. Villani, *Optimal Transport: Old and New* (Springer, 2009).

[16] L. Dammeier, R. Schwonnek, and R. Werner, "Uncertainty relations for angular momentum," New J. Phys. **17**, 093046 (2015).

[17] P. Lévy, "L' addition des variables aléatoires définies sur une circonférence," Bull. Soc. Math. France **67**, 1–41 (1939).

[18] R. von Mises, "Über die 'Ganzzahligkeit' der Atomgewichte und verwandte Fragen," Phys. Z. **19**, 490 (1918).

[19] E. Breitenberger, "Uncertainty measures and uncertainty relations for angle observables," Found. Phys. **15**, 353 (1985).

[20] M. B. M.A. Alonso, "Mapping-based width measures and uncertainty relations for periodic functions," Signal Processing **84**, 2425 (2004).

[21] R. Werner, "Quantum harmonic analysis on phase space," J. Math. Phys. **25**, 1404 (1984).

[22] N. Dunford and J. Schwartz, *Linear Operators, Part I: General theory* (Wiley, 1957).

[23] M. Abramowitz and I. Stegun, *Handbook of Mathematical Functions* (Dover Publications, 1965).

[24] J. Řeháček, Z. Bouchal, R. Čelechovský, Z. Hradil, and L. Sánchez-Soto, "Experimental test of uncertainty relations for quantum mechanics on a circle," Phys. Rev. A **77**, 032110 (2008).

[25] M. Reed and B. Simon, *Methods of Modern Mathematical Physics, Vol. IV: Analysis of Operators* (Academic Press, 1978).

[26] D. Judge, "On the uncertainty relation for angle variables," Nuovo Cimento **31**, 332 (1964).

[27] M. Bouten, N. Maene, and P. V. Leuven, "On an uncertainty relation for angular variables," Nuovo Cimento **37**, 1119 (1965).

[28] M. Nieto, "Quantum phase and quantum phase operators: some physics and some history," Physica Scripta **T48**, 5 (1993).

[29] M. N. P. Carruthers, "Phase and angle variables in quantum mechanics," Rev. Mod. Phys. **40**, 411 (1968).

[30] R. Jackiw, "Minimum uncertainty product, number-phase uncertainty product, and coherent states," J. Math. Phys. **9**, 339 (1968).

[31] P. Lahti, J.-P. Pellonpää, and J. Schultz, "Number and phase: complementarity and joint measurement uncertainties," Journal of Physics A: Mathematical and Theoretical **50**, 375301 (2017).