Uncertainty Relations for Positive Operator Valued Measures

Serge Massar

Laboratoire d’Information Quantique, C.P. 225, Université Libre de Bruxelles (U.L.B.), Boulevard du Triomphe, B-1050 Bruxelles, Belgium

(Dated: June 9, 2021)

How much unavoidable randomness is generated by a Positive Operator Valued Measure (POVM)? We address this question using two complementary approaches. First we study the variance of a real variable associated to the POVM outcomes. In this context we introduce an uncertainty operator which measures how much additional noise is introduced by carrying out a POVM rather than a von Neumann measurement. We illustrate this first approach by studying the variances of joint estimates of $\sigma_x$ and $\sigma_z$ for spin 1/2 particles. We show that for unbiased measurements the sum of these variances is lower bounded by 1. In our second approach we study the entropy of the POVM outcomes. In particular we try to establish lower bounds on the entropy of the POVM outcomes.

We illustrate this second approach by examples.

I. INTRODUCTION

Richard Feynman famously said that “no one understands quantum mechanics”. Nevertheless, practitioners of quantum mechanics have a number of mental pictures that help them reason about the theory. Arguably the most important of them is the uncertainty principle [1], which both gives fundamental insights and provides important quantitative predictions.

The uncertainty principle itself comes in several forms. The best known bounds the variances of two operators $A = \sum_k a_k|a_k\rangle\langle a_k|$ and $B = \sum_l b_l|b_l\rangle\langle b_l|$ measured in the quantum state $|\psi\rangle$. The resulting inequality [2] is:

$$\Delta A \Delta B \geq \frac{1}{2} |\langle \psi | [A, B] |\psi\rangle|$$

(1)

where $\Delta A^2 = \langle A^2 \rangle - \langle A \rangle^2$, $\Delta B^2 = \langle B^2 \rangle - \langle B \rangle^2$. In terms of position and momentum operators it takes the form [3]:

$$\Delta x \Delta p \geq \frac{\hbar}{2}$$

(2)

An alternative formulation based on entropic uncertainty relations [4] provides bounds on the entropy of the outcomes of measurements of $A$ and $B$. The version conjectured in [5] and proved in [6] (for the particular case of position and momentum see [7]) takes the form:

$$H(A) + H(B) \geq -\log_2 \max_{kl} |\langle a_k | b_l \rangle|^2$$

(3)

where $H(A) = -\sum_k p_A(k) \log_2 p_A(k)$ is the Shannon entropy of the probability distribution $p_A(k) = |\langle a_k |\psi\rangle|^2$ and similarly for $H(B)$. Entropic uncertainty relations have some conceptual advantages over (1): they are independent of the values $a_k$, $b_l$ one assigns (often arbitrarily) to the outcomes of the measurements of $A$ and $B$; and their right hand side (rhs) is independent of the quantum state $|\psi\rangle$, whereas the rhs of (1) can vanish even though $\Delta A$ and $\Delta B$ are both positive. Note that in the above formulations the measurements of $A$ and $B$ are mutually exclusive: they cannot both be carried out.

A conceptually different form of the uncertainty principle concerns measurements that simultaneously estimate two non commuting operators $A$ and $B$. The precision $\Delta A$ and $\Delta B$ with which $A$ and $B$ are jointly estimated should obey a constraint similar to eq. (1). But proving such constraints is difficult, see e.g. [8, 9, 10, 11, 12, 13, 14, 15, 16, 17].

Simultaneously estimating two observables cannot be carried out within the usual framework of Projector Valued Measures (PVM). Rather it must be formulated within the more general context of Positive Operator Valued Measures (POVM). Formally a POVM $M$ is described by a set of positive operators that sum to the identity: $M = \{m_k\}$, $m_k \geq 0$, $\sum_k m_k = I$.

POVM’s play an essential role in measurement theory as they describe measurements affected by noise, fuzzy measurements, measurements that simultaneously estimate two observables, etc... POVM’s also play an essential role in quantum information: they are often the measurements which allow the most information to be extracted from a quantum system [19, 20], they are widely used for quantum communication tasks, etc... However POVM’s suffer from the fact that (except when some of the POVM elements are projectors) there is no state $|\psi\rangle$ for which the outcome of the POVM is fixed. That is, the outcomes of POVM’s are affected by unavoidable noise. This raises a fundamental question: why is it that, given their unavoidable noise, POVM’s are nevertheless sometimes better than PVM’s for information processing tasks? A partial explanation is that the noise of a POVM is more uniformly distributed over Hilbert space: there are many states on which a PVM will give completely random results, whereas many POVM’s never give completely random outcomes. But it is unclear whether this is the whole answer. Indeed there at least one very specific context in which the randomness produced by a POVM can be removed: if the same POVM must be carried out on many independent states, then one can devise a collective POVM acting on all the states which is almost a PVM, and which when restricted to a single system, acts as the original POVM [21].

The inherent randomness of POVM’s may also explain
some of their limitations. For instance there is to our knowledge no non locality experiment for which POVM's are better than PVM’s, and this is probably due to the added noise coming from POVM’s. Better understanding the potentialities and limitations of POVM’s thus hinges on better understanding the unavoidable noise they add to experiments. In the present work we address this task, and carry out a systematic study of the noise produced by POVM’s. There are already many works that address aspects of this question [8, 9, 10, 11, 12, 13, 14, 15], often in the context of joint measurements of position and momentum. But it is clear that a unified approach is necessary that does not focus on the technical difficulties of infinite dimensional spaces, but rather on the conceptual issues involved. The present work aims to fill this gap by following two complementary approaches, similar in spirit to the complementary approaches provided by the Robertson and entropic uncertainty relations mentioned above.

II. UNCERTAINTY OPERATOR

In our first approach (see also [8, 9, 12, 16, 17]) we suppose that to each POVM element \( m_k \) one associates a real value \( \mu_k \in \mathbb{R} \). This association of course suffers from the same limitations as the Robertson inequality: the choice of the \( \mu_k \) is to some extent arbitrary. Different choices will yield different estimates of the uncertainty. Furthermore, it is often natural to associate several values to the same POVM element \( m_k \). We will then have several uncertainties associated with the same POVM.

The expectation value of \( \mu_k \) is

\[
\bar{\mu} = \sum_k \mu_k \langle \psi | m_k | \psi \rangle = \langle \psi | M | \psi \rangle \tag{4}
\]

where we introduce the operator

\[
M = \sum_k \mu_k m_k . \tag{5}
\]

The variance of \( \mu_k \) is

\[
\text{Var}(\mu) = \sum_k \mu_k^2 \langle \psi | m_k | \psi \rangle - \bar{\mu}^2 \tag{6}
\]

\[
= \langle \psi | M^2 - \bar{\mu}^2 | \psi \rangle + \langle \psi | \Delta M^2 | \psi \rangle \tag{7}
\]

where we introduce the uncertainty operator

\[
\Delta M^2 = \sum_k \mu_k^2 m_k - M^2 \tag{8}
\]

The first term in eq. (7) is the variance of \( \mu_k \) which would arise if one was to “measure the observable \( M \)” in the usual sense. Note that there are always some states \( |\psi\rangle \) for which this term vanishes. The second term is the additional uncertainty which arises because one is measuring a POVM and not a PVM. It does not depend on the average value \( \bar{\mu} \). The uncertainty operator \( \Delta M^2 \) thus characterizes the extra noise coming from the POVM. It has the following important properties:

1. Positivity: \( \Delta M^2 \geq 0 \) is a positive operator, which follows from rewriting it in the manifestly positive form \( \Delta M^2 = \sum_k (M - \mu_k) m_k (M - \mu_k) \tag{16} \):

2. Vanishing on PVM’s: \( \Delta M^2 = 0 \) if \( M \) is a PVM;

3. Strict positivity on POVM’s: if \( M \) is not a PVM, then there exists a choice of \( \mu_k \) such that \( \Delta M^2 > 0 \) is strictly positive (simply take \( \mu_k = \delta_{kk_0} \) to be zero except for one value \( k_0 \), with \( k_0 \) such that \( m_{k_0} \) is not a projector);

4. Reduction to classical random variables: if the POVM elements are all proportional to the identity \( m_k = |m_k| \mathbb{I} \), then the probabilities of the outcomes are independent of the quantum state, and \( \mu_k \) is a classical random variable, with associated probabilities \( p_k = |m_k|/d \) where \( d \) is the dimension of the Hilbert space. In this case the first term in eq. (4) vanishes and \( \Delta M^2 = \mathbb{I} \sum_k (\mu_k - \bar{\mu})^2 p_k \) is the variance of the classical random variable \( \mu_k \).

5. Additivity under tensor product: Consider two POVM’s and their associated values \( \{m_i, \mu_i\} \) and \( \{n_j, \nu_j\} \) acting on different systems. We construct the tensor product POVM as \( \{m_i \otimes n_j, \mu_i + \nu_j\} \), where we associate to each outcome \((i,j)\) the sum of the values \( \mu_i + \nu_j \) for each outcome. Then the uncertainty operator for the tensor product POVM is \( \Delta(M \otimes N)^2 = \Delta M^2 \otimes \mathbb{I} + \mathbb{I} \otimes \Delta N^2 \).

6. Increase under convex combination: Consider two POVM’s and their associated values \( \{m_i, \mu_i\} \) and \( \{n_j, \nu_j\} \) acting on the same system. We construct the convex combination of these two POVM’s by realizing the first POVM with probability \( p \), and the second POVM with probability \( q \) \((p + q = 1)\) to obtain a POVM with elements \( pm_i + qn_j \) to which are associated the values \( \mu_i \) and \( \nu_j \) respectively. The uncertainty operator for the convex combination POVM is \( \Delta(pM + qN)^2 = p\Delta M^2 + q\Delta N^2 + pq(M - N)^2 \).

Properties 5 and 6 are in direct analogy with the way the variances of independent classical random variables behave under addition and convex combination.

Following the ideas developed in [16], see also [17], some further properties of the uncertainty operator can be obtained by introducing Naimark’s extension of the POVM \( M \). Naimark’s theorem states that there exists an extended Hilbert space \( \tilde{H} \) which is the direct sum of
the system Hilbert space $H$ (on which $\mathcal{M}$ acts) and an ancillary Hilbert space $H'$ ($H = H \oplus H'$); and a basis of the extended Hilbert space $|\tilde{m}_k\rangle$ ($\langle \tilde{m}_k | \tilde{m}_k' \rangle = \delta_{k,k'}$) which restricted to the system space gives the POVM $\mathcal{M}$: $|\tilde{m}_k\rangle = |m_k\rangle + |m_k'\rangle$, where $|m_k\rangle \in H$ and $|m_k'\rangle \in H'$. (For simplicity we suppose that the POVM elements are rank 1: $m_k = |m_k\rangle \langle m_k|$, the general case follows by taking several of the $\mu_k$ to have the same value).

We can now then define a Hermitian operator on the extended space as $M_E = \sum_k \mu_k |\tilde{m}_k\rangle \langle \tilde{m}_k|$. Then, if we denote by $P$ the projector onto the system space $H$, we can rewrite the uncertainty operator as $\Delta M^2 = P(M_E - M)^2 P$ which shows that the uncertainty operator measures the difference between $M$ and any Naimark extension of the POVM $\mathcal{M}$. More generally, it is possible to introduce a notion of distance between the POVM $\mathcal{M}$ and Hermitian operators $A$ acting on $H$ by $d(A,M)^2 = \text{tr}[P(M_E - A)^2 P]$. The Hermitian operator which minimises this distance is $M$, and the minimum distance is $d(M,M)^2 = \text{tr}\Delta M^2$. We can also consider the statistical distance between $M_E$ and Hermitian operators $A$ acting on $H$: $|\psi \rangle (M_E - A)^2 |\psi \rangle$ where $|\psi \rangle \in H$ is a state in the system Hilbert space. Once more the statistical distance is minimised if we take $A = M$, and the minimal statistical distance is $\langle \psi | \Delta M^2 | \psi \rangle$.

Finally it is interesting to note that the uncertainty relation eq. (7) can also be derived from the Robertson inequality.

We shall now generalise eq. (11) to arbitrary POVM’s eq. (9), see [13] and references therein. Another illustration of the power of our approach is provided by the problem of jointly estimating both the $\sigma_x$ and $\sigma_z$ observables of a spin 1/2 particle. Such a measurement will be described by a POVM with 4 outcomes: $m_{zx}$, where the label $z = \pm 1$ ($x = \pm 1$) corresponds to inferring that a measurement of $\sigma_z$ ($\sigma_x$) would preferentially have given the value $z = +1$ or $z = -1$ ($x = +1$ or $x = -1$) outcome.

Let us first illustrate such a POVM by the following example:

$$m_{zx} = \frac{1}{4} + \frac{z \cos \theta}{4} \sigma_z + \frac{z \sin \theta}{4} \sigma_x.$$  

(9)

If we use this POVM to estimate the $z$ ($x$) component of the spin, then the associated operators are $Z = \sum_{z,x = \pm 1} zm_{zx} = \cos \theta \sigma_z$ and $X = \sum_{z,x = \pm 1} xm_{zx} = \sin \theta \sigma_x$. Thus one is indeed simultaneously estimating the $z$ and $x$ components of the spin, but with reduced sensitivity with respect to measuring the observables $\sigma_z$ and $\sigma_x$ separately. The associated uncertainty operators are

$$\Delta Z^2 = \sum_{zx} z^2 m_{zx} - (\sum_{zx} zm_{zx})^2 = (1 - \cos^2 \theta) \mathbb{1}$$

$$\Delta X^2 = \sum_{zx} x^2 m_{zx} - (\sum_{zx} xm_{zx})^2 = (1 - \sin^2 \theta) \mathbb{1}$$

(10)

which we can group in the relation $\Delta X^2 + \Delta Z^2 = 1$ which implies the constraint on the sum of the variances

$$\text{Var}(Z) + \text{Var}(X) \geq 1.$$  

(11)

This constraint has been derived previously for the particular POVM’s eq. (9), see [13] and references therein.

Let us now generalise eq. (11) to arbitrary POVM’s with 4 outcomes $z,x = \pm 1$, where outcome $z = +1, x = +1$ corresponds to guessing that measurements of $\sigma_z$ and $\sigma_x$ would both have given outcome +1; and similarly for the other values of $z,x$. We require that the POVM be unbiased in the following weak sense: we require that for the completely mixed state $\rho = \mathbb{1}/2$, the probabilities of the different outcomes $z,x$ are all equally probable.

B. Joint Estimates of $\sigma_z$ and $\sigma_z$

III. APPLICATIONS

A. Joint Measurements of Position and Momentum

A particularly well studied example (for more details see [12] and references therein) is the covariant joint measurement of $p$ and $x$ of the form $m_{px} = D_{px} m_0 D_{px}$ where $D_{px}$ is the displacement operator in phase space and $m_0$ is a normalized state.

If one uses $m_{pq}$ to estimate $q$, then one will associate with $m_{pq}$ the value $q$. In this case one finds that the associated operator $Q = \int dp dq \; q |m_0\rangle \langle q|$ is just the operator $q$ displaced to coincide with the center of $m_0$, and the uncertainty operator is just the identity times the variance of $q$ in the state $m_0$: $\Delta Q^2 = \mathbb{1} \Delta q_{m_0}^2$ with $\Delta q_{m_0}^2 = \left( \int dq dq \langle q | m_0(q) \rangle - \langle q | m_0 | q \rangle \right)^2$. From eq. (7) it then follows that the variance of this estimator of $q$ in state $|\psi \rangle$ is the sum of the variances of $q$ in states $|\psi \rangle$ and $m_0$: $\Delta q^2 = \Delta q_{m_0}^2 + \Delta q_{m_0}^2$. Similarly, if one associates with $m_{pq}$ the value of $p$, one finds that the variance of this estimator of $p$ in state $|\psi \rangle$ is $\Delta p^2 = \Delta p_{m_0}^2 + \Delta p_{m_0}^2$. Putting these two results together and using eq. (9), we obtain a bound on the variances of the joint estimates $p$ and $q$: $\Delta q^2 \Delta p^2 \geq \Delta q_{m_0}^2 \Delta p_{m_0}^2 \geq \hbar^2$ which is twice the variance of $q$ and $p$ separately.
This implies that the POVM elements can be written as
\[ m_{zx} = \frac{1}{4} + \bar{v}_{xz} \cdot \bar{\sigma} \] with \( |\bar{v}_{xz}| \leq 1/4 \) and \( \sum_{zx} \bar{v}_{xz} = 0 \). In addition for the measurement to be faithfull we require that \( \bar{v}_{+z} \) has positive \( z \) and \( x \) components. This corresponds to requiring that the state \( \psi \) that maximizes the probability of getting outcome ++ would give with high probability \( \sigma_z = +1 \) and \( \sigma_x = +1 \) if one were to measure these operators. Similar conditions hold for the other values of \( z \) and \( x \). As before we associate with outcome \( z, x \) the values \( Z = z, X = x \). It then follows (in fact from the unbiasedness condition alone) that
\[ \Delta Z^2 + \Delta X^2 = 1 \] which implies eq. (11). Note also that equality in eq. (11) is obtained if and only if \( \bar{v}_{+z} = -\bar{v}_{-z}, \bar{v}_{+x} = -\bar{v}_{-x} \) and \( |\bar{v}_{xz}| = 1/4 \). Thus the class of unbiased POVM which jointly estimates \( \sigma_z \) and \( \sigma_x \) with the least added noise contains, but is larger, than the example eq. (10). (Note that all these optimal POVM’s can be realized by measuring with probability 1/2 the operator \( \bar{v}_{-z} \cdot \bar{\sigma} \) and with probability 1/2 the operator \( \bar{v}_{+z} \cdot \bar{\sigma} \).

The above problem of jointly measuring \( \sigma_z \) and \( \sigma_x \) for a spin 1/2 particle is related to, and motivated by, the problem of better understanding uncertainty relations for Mach-Zehnder interferometers. Indeed in the case of a Mach-Zehnder interferometer there are two natural questions to ask: through which arm of the interferometer does the particle pass, and by which output port does it exit. Because the Mach-Zehnder interferometer is a two dimensional system, there is a formal mapping to the spin 1/2 case, with the path observable associated to \( \sigma_z \) and the output port observable associated with \( \sigma_x \).

Two uncertainty relations have been derived for the Mach-Zehnder problem. The first is similar in spirit to the Robertson inequality, as it relates the variances of two incompatible observables measured on the same state \( \psi \). It relates the Predictability \( P \) of the path taken by the particle to the Visibility \( V \) of the interference fringes by \( P + V \leq \frac{1}{2} \) (see also [22, 23]). We have the following correspondence with the spin 1/2 case: \( P = 1 - \text{Var}(\sigma_z)_\psi \) and \( V \geq 1 - \text{Var}(\sigma_x)_\psi \). (The inequality in the last relation comes from the fact that the visibility is defined by the maximum contrast when the phase in the interferometer is varied. The maximum contrast may not occur for a measurement of \( \sigma_x \), but of \( \cos \theta \sigma_x + \sin \theta \sigma_y \) for some \( \theta \).

The second uncertainty relation concerns \textit{joint estimates} of the path a particle takes in a Mach-Zehnder interferometer and the output port by which it exits (see also [13, 18]). The approach adopted in these works is to attach to the particle a which way marker which registers information about which path the particle took in the interferometer. One then measures both the marker and the output port, thereby implementing a POVM with four outputs (see [22, 23, 24] for experimental realisations of this idea). Englert [14] obtained in this context an uncertainty relation which constrains the likelihood of correctly guessing the path taken by the particle (the distinguishability \( D \)) and the visibility \( V \) of the interference fringes) in a joint measurement using a which way marker: \( D + V \leq 1 \).

Our result eq. (11) is related to the work of Englert, since it constrains the precision with which one can \textit{jointly estimate} the path and output port taken by the particle. But here we have adopted a more abstract approach, and consider arbitrary unbiased POVM’s with 4 outcomes, rather than specific implementations using which way markers.

\section{IV. Entropic Uncertainty Relation for POVM’s}

An alternative approach toward understanding the amount of randomness generated by a POVM is to lower bound entropy of the outcomes \( H(M) = - \sum_k p(k) \log_2 p(k) \) with \( p(k) = \langle \psi | m_k | \psi \rangle \). A very simple bound of this type is provided by the largest eigenvalue of the POVM elements \( m_k \):\[ H(M) \geq - \log_2 \left( \max_k \langle \psi | m_k | \psi \rangle \right) . \] (13)

Thus, except in the trivial case when one of the POVM elements is a projector, the entropy of the POVM outcomes is always positive. In some cases we have been able to improve this bound.

When the POVM can be realized by carrying out with probability 1/2 one of two non degenerate PVM’s \( \{ |a_k \rangle \} \) and \( \{ |b_l \rangle \} \) (with \( \langle a_k | a_{k'} \rangle = \delta_{k k'} \) and \( \langle b_l | b_{l'} \rangle = \delta_{l l'} \)) then eq. (13) implies \( H(M) \geq 1 - \frac{1}{2} \log_2 \text{max}_{k,l} \langle a_k | b_l \rangle \). Applied to the POVM eq. (9) this yields the bound
\[ H(\{m_{zx}\}) \geq 1 - \log_2 \cos \theta , \quad 0 \leq \theta \leq \pi/4 \] (14)
(which is tight when \( \theta = 0 \) or \( \theta = \pi/4 \) but is suboptimal in between). From eq. (14) it follows that as the POVM \( \{m_{zx}\} \) goes from estimating only \( \sigma_z \) (\( \theta = 0 \)) to estimating \( \sigma_z \) and \( \sigma_x \) with equal sensitivity (\( \theta = \pi/4 \)), the minimum entropy generated by the POVM increases from 1 to 3/2 bits. This is a particular example of what we expect is a general trade off: POVM’s that probe more uniformly the Hilbert space (which can be a useful property for information processing) generate more randomness (a deleterious property).

We have also improved on eq. (13) using a different method. Let us consider bounds on the entropies of two POVM’s, \( M = \{ |m_k \rangle \langle m_k | \} \) and \( N = \{ |n_l \rangle \langle n_l | \} \), whose elements are all rank 1, acting on the same state. As was noted in [27], the proof of the entropic uncertainty relation given in [6] immediately generalizes to:
\[ H(M) + H(N) \geq - \log_2 \text{max}_{k,l} \langle m_k | n_l \rangle . \] (15)

This can be strengthened by noting that the Naimark extension is not unique: any POVM of the form \( U' | m_k \rangle = | m_k \rangle + U' | m'_k \rangle \) with \( U' \) acting only on the ancillary Hilbert
space $H'$ is a possible extension of the POVM $\mathcal{M}$. Applying eq. (3) to the Naimark extension of the POVM's $\mathcal{M}$ and $\mathcal{N}$ and taking the best such bound yields:

$$H(M) + H(N) \geq \max_{U'} \log_2 \max_{kl} |\tilde{m}_k(U')|.$$  \hfill (16)$$

We can now go back to the case of a single POVM. Let us take in eq. (16) the two POVM's to be identical but with different Naimark extensions. This yields the entropic uncertainty relation for a single POVM:

$$H(M) \geq \max_{U'} \frac{1}{2} \log_2 \max_{kl} |\tilde{m}_k(U')|.$$  \hfill (17)$$

As an illustration we consider the POVM described in [20] which acts on the symmetric space of two spin 1/2 particles and is composed of 4 elements, each proportional to the projector onto two parallel spins oriented along the 4 corners of a tetrahedron. The Naimark extension of this POVM can be constructed explicitly as $|\tilde{m}_{(a)}\rangle = \frac{1}{\sqrt{2}}(|\uparrow_z\rangle + |\downarrow_z\rangle) + \frac{1}{\sqrt{2}}(|\uparrow_j\rangle + |\downarrow_j\rangle)$ where $|\uparrow_j\rangle = (|\uparrow_z\rangle + \sqrt{2}/2)_{j=\{1,2,3\}}$. If we take $U'|a\rangle = -|a\rangle$ (with $U'$ acting as the identity on the space of the two spins), then eq. (17) implies that the entropy of this POVM is bounded by $H(M) \geq 1$ bit, which is significantly better than the bound $H(M) \geq -\log_2(3/4)$ which follows from eq. (13).

V. CONCLUSION

In summary, in this paper we have studied how much unavoidable randomness is generated by a Positive Operator Valued Measure (POVM). We address this question using two complementary approaches. First we study the variance of a variable associated to the POVM outcomes. We illustrate this method by establishing an inequality on the variances of joint estimates of $\sigma_x$ and $\sigma_z$ for spin 1/2 particles. Second we study lower bounds on the entropy of the POVM outcomes. We also illustrate this second approach by examples.

Acknowledgements I would like to thank Jeremy O’Brien, Paul Bush, and specially Michael Hall, for comments on the first version of this manuscript. I acknowledge support by EU project QAP contract 015848, and by the IAP project -Belgium Science Policy- P6/10 Photonics@be.