EPJ Web of Conferences **78**, 01002 (2014)
DOI: 10.1051/epjconf/20147801002

Quantum rms error and Heisenberg's error-disturbance relation

Paul Busch[a]

Department of Mathematics, University of York, York, UK

Abstract. Reports on experiments recently performed in Vienna [Erhard *et al*, *Nature Phys.* **8**, 185 (2012)] and Toronto [Rozema *et al*, *Phys. Rev. Lett.* **109**, 100404 (2012)] include claims of a violation of Heisenberg's error-disturbance relation. In contrast, a Heisenberg-type tradeoff relation for joint measurements of position and momentum has been formulated and proven in [*Phys. Rev. Lett.* **111**, 160405 (2013)]. Here I show how the apparent conflict is resolved by a careful consideration of the quantum generalisation of the notion of root-mean-square error. The claim of a violation of Heisenberg's principle is untenable as it is based on a historically wrong attribution of an incorrect relation to Heisenberg, which is in fact trivially violated. We review a new general trade-off relation for the necessary errors in approximate joint measurements of incompatible qubit observables that is in the spirit of Heisenberg's intuitions. The experiments mentioned may directly be used to test this new error inequality.

1 Introduction

Heisenberg's uncertainty principle, as it was conceived in his classic work of 1927 [1], may be understood as comprising three distinct statements concerning pairs of canonically conjugate quantities such as the position and momentum observables of a quantum particle:

Preparation Uncertainty: position and momentum distributions cannot both be arbitrarily sharply concentrated in the same state.

Measurement Uncertainty: the errors in any joint measurement of position and momentum cannot both be arbitrarily small.

Error-Disturbance Relation: the error in any position (momentum) measurement and the disturbance of momentum (position) caused by this measurement cannot both be arbitrarily small.

Heisenberg's focus was on joint measurement uncertainty and the error-disturbance tradeoffs but he deduced these by taking recourse to preparation uncertainty. He expressed the error-disturbance relation by the symbolic relation

$$p_1\, q_1 \sim h, \tag{1}$$

where q_1 stands for, say, the error in a position measurement and p_1 the ensuing disturbance of momentum; the necessary order of magnitude for the product is given by Planck's constant h.

From a modern perspective it seems appropriate to capture the spirit of Heisenberg's principle as the general statement that the errors and disturbances in a joint measurement of two observables are

[a]e-mail: paul.busch@york.ac.uk

constrained by the degree of their incompatibility. It has taken many decades until serious attempts started at giving rigorous formulations and derivations of such measurement uncertainty relations as consequences of quantum mechanics. The reason for this must be seen in the fact that until recently the tools of quantum measurement theory were not sufficiently developed to enable a conceptualisation of measurement errors and disturbances.

In a series of papers starting around 2002 [2–7], Ozawa presented a rigorous formulation of an inequality that he considered to express such an error-disturbance trade-off. Ozawa's inequality suggests that a naive generalisation of Heisenberg's intuitive relation (1) is bound to fail. When in 2012 experimental confirmations of Ozawa's inequality were published, the claim of a violation of Heisenberg's principle made in these papers [8, 9] stirred up a considerable media hype. Here we review an analysis of the problem of quantifying measurement errors in quantum mechanics and present new measurement uncertainty relations for qubits that render the claim untenable. The present contribution is a concise report on joint work with P. Lahti and R.F. Werner [10–13].

2 Noise operators: a failed quantum rendering of rms error

Ozawa's approach is based on an adaptation of the concept of amplifier noise to the context of quantum measurement. The proposed measure of error, $\varepsilon_{\rm NO}$, is given, rather suggestively at first sight, as the root of the expectation of the squared difference of the target observable to be approximated and the read-out operator actually measured. This difference of operators is known as the *noise operator* in quantum optics and taken by Ozawa here to represent "measurement noise". Thus, let A denote the target observable and $\mathcal{M}$ a measurement scheme comprising a unitary coupling U between the object system and probe system, the density operator σ of the probe, and the read-out operator Z. Then according to Ozawa the error and disturbance appropriate to any input state ρ of the object are:[1]

$$\varepsilon_{\rm NO}(A,\mathcal{M},\rho) = \left\langle (U^*\mathbf{1}\otimes ZU - A\otimes\mathbf{1})^2\right\rangle_{\rho\otimes\sigma}^{1/2}, \quad \eta_{\rm NO}(B,\mathcal{M},\rho) = \left\langle (U^*B\otimes\mathbf{1}U - B\otimes\mathbf{1})^2\right\rangle_{\rho\otimes\sigma}^{1/2}. \tag{2}$$

Ozawa then proceeds to point out (correctly) that the inequality

$$\varepsilon_{\rm NO}(A,\mathcal{M},\rho)\eta_{\rm NO}(B,\mathcal{M},\rho) \geq \tfrac{1}{2}\left|\langle[A,B]\rangle_\rho\right| \tag{3}$$

is not generally valid but may be violated for any pair of noncommuting operators A, B. Since he (wrongly) attributes (3) to Heisenberg, he is then able to claim a violation of Heisenberg's error-disturbance relation. His modification and correction of (3) has become known as Ozawa's inequality:

$$\varepsilon_{\rm NO}(A,\mathcal{M},\rho)\eta_{\rm NO}(B,\mathcal{M},\rho) + \varepsilon_{\rm NO}(A,\mathcal{M},\rho)\Delta(B,\rho) + \Delta(A,\rho)\eta_{\rm NO}(B,\mathcal{M},\rho) \geq \tfrac{1}{2}\left|\langle[A,B]\rangle_\rho\right|; \tag{4}$$

here $\Delta(A,\rho)$ and $\Delta(B,\rho)$ denote the standard deviations of A and B, respectively, in the state ρ.

We emphasise that the attribution of (3) to Heisenberg is historically unjustified since Heisenberg never made a statement of this precision or generality. The closest he comes in [1] to formally defining measurement error is by relating it to the standard deviation in the output state of an (approximately) repeatable measurement. In his Chicago lectures [14] he explicitly endorses Kennard's [15] formulation of the preparation uncertainty relation in terms of standard deviations as the basis also for measurement uncertainty and error-disturbance relations.

[1] We use the shorthand $\langle X\rangle_\rho$ for expectation values $\mathrm{tr}[X\rho]$.

2.1 The failure of $\varepsilon_{\rm NO}, \eta_{\rm NO}$ as measures of error and disturbance

Disregarding the wrong attribution to Heisenberg, another issue with (3) and likewise with (4) is that the interpretation of these inequalities as error-disturbance trade-off relations hinges on the question whether the quantities $\varepsilon_{\rm NO}, \eta_{\rm NO}$ are in fact faithful measures and indicators of error and disturbance. As shown in a detailed analysis in [10], this question must be answered in the negative.

For any quantity to represent a measure of error (and similarly for disturbance), it is reasonable to require that the quantity is indicative of the difference of the *distributions* of the two observables being compared in the error analysis. We call a quantity *operationally significant* as an error measure if it has this property. The operators appearing in (2) are differences of two operators that may not commute, in general, and in such cases there is no guarantee that the expectation values $\varepsilon_{\rm NO}, \eta_{\rm NO}$ are operationally significant. It is, in fact, easy to give examples showing their failure in both directions: for example, $\varepsilon_{\rm NO}$ can be zero where the distributions are different; or it may be nonzero where the distributions are identical.

Example 1 Consider an attempted approximate measurement of the position Q of a particle by means of an observable given by

$$Q' = Q + \alpha\left(\frac{P^2}{2m} + \frac{m\omega^2}{2}Q^2 - \frac{1}{2}\hbar\omega\mathbf{1}\right), \quad \alpha > 0.$$

Then Q' is a discrete-spectrum operator, and $\varepsilon_{\rm NO}(Q, Q', \psi)$ is determined as

$$\varepsilon_{\rm NO}(Q, Q', \psi)^2 = \langle\psi|(Q' - Q)^2\psi\rangle.$$

This vanishes if ψ is the harmonic oscillator ground state, while the distributions of Q and Q' are quite different, one being continuous, the other discrete.

Example 2 It turns out that Ozawa's inequality undermines its own intended interpretation as an error-disturbance relation. To see this, consider an accurate measurement $\mathcal{M}$ of A. In this case $\varepsilon_{\rm NO}(A, \mathcal{M}, \rho) = 0$, and thus inequality (4) reduces to

$$\eta_{\rm NO}(B, \mathcal{M}, \rho) \geq \frac{|\langle[A, B]\rangle_\rho|}{2\Delta(A, \rho)}. \tag{5}$$

If A and B do not commute in ρ then according to Ozawa's inequality, the disturbance of observable B should be nonzero, suggesting that the distribution of B before and after the A measurement must differ. However, the measurement $\mathcal{M}$ could be chosen such that its nonselective state transformation is a constant channel, $\rho \mapsto \rho_0$; then, if $\rho = \rho_0$, the measurement has caused no disturbance whatsoever, and the distribution of B has not been changed through the measurement. Hence, it must be concluded, the non-vanishing $\eta_{\rm NO}$ cannot represent disturbance. In fact, closer analysis of this example shows that $\eta_{\rm NO}$ contains contributions arising from preparation uncertainty and is thus not a faithful measure of disturbance, and (5) simply turns into the usual preparation uncertainty relation.

Example 3 One can also consider a trivially nondisturbing measurement scheme with vanishing $\eta_{\rm NO}$: let U be the identity, then all states of the object remain unchanged, and $\eta_{\rm NO}(B, \mathcal{M}, \rho) = 0$ for all ρ and all B. The measured observable is trivial, giving the same output distribution E^Z_σ for every state ρ. (Here E^Z denotes the spectral measure of the pointer operator Z and E^Z_σ the associated probability distribution in the probe state σ.) If the probe is chosen to be a copy of the object system and the pointer $Z = A$, then for the state $\rho = \sigma$, the output distribution $E^Z_\sigma = E^A_\rho$, which is to say that the distribution of A is accurately reproduced in this particular state. (A broken clock shows the correct time exactly twice a day.) By contrast, Ozawa's inequality entails

$$\varepsilon_{\rm NO}(A,\mathcal{M},\sigma) \geq \frac{|\langle [A,B]\rangle_\sigma|}{2\Delta(B,\sigma)},$$

where it is always possible to choose B,σ such that the lower bound is nonzero. Hence $\varepsilon_{\rm NO}$ does not represent measurement error correctly. Again, this is (at least partly) due to the presence of preparation uncertainty contributions in $\varepsilon_{\rm NO}$.

There are cases of interest where $\varepsilon_{\rm NO}$ does have some operational significance: the case where the target observable A and the actually measured observable C commute (in the given state); and the case of unbiased measurements. To discuss these situations we note that the quantity $\varepsilon_{\rm NO}(A,\mathcal{M},\rho)$ can be expressed in terms of the first and second moments $C[x], C[x^2]$ of the positive operator valued measure (POVM) C measured by $\mathcal{M}$; alternatively, one can write it in a form that looks superficially like a mean squared deviation, except that the bimeasure is not a probability bimeasure if the observables A and C do not commute:

$$\varepsilon_{NO}(A,C,\rho)^2 = \mathrm{tr}\left[\rho(C[x]-A)^2\right] + \mathrm{tr}\left[\rho(C[x^2]-C[x]^2)\right] \tag{6}$$

$$= \iint (x'-x)^2\,\mathrm{Re}\,\mathrm{tr}\left[\rho E^A(dx)C(dx')\right]. \tag{7}$$

2.2 The commutative case

Another condition under which $\varepsilon_{\rm NO}, \eta_{\rm NO}$ are operationally significant is the commutativity of the observables being compared. In fact, if A and C commute (as POVMs) in the state ρ, then they have a joint probability distribution given by $\mathrm{tr}\left[\rho E^A(dx)C(dx')\right]$; in this case the formula (7) is operationally well defined, giving the root of the mean squared deviation between the simultaneously obtained values of A and C in the joint probability distribution $\mathrm{tr}\left[\rho E^A(dx)C(dx')\right]$.

It is here where the meaning and limitation of application of $\varepsilon_{\rm NO}$ (and similarly of $\eta_{\rm NO}$) becomes apparent: these quantities are expectation values for joint measurements of the target and approximator observables A, C. Such measurements yield simultaneous values for A (representing the "true" value) and for C, the approximate value, and $\varepsilon_{\rm NO}$ quantifies the rms deviation of the pairs of values obtained one by one, that is, $\varepsilon_{\rm NO}$ represents a *value-comparison* method for error estimation. This interpretation is only available when A and C can be measured jointly in the given state, which is a rather exceptional situation. In cases where A and C are not jointly measurable, one can only compare the A and C distributions, hence the error estimation is based on a *distribution comparison*. This latter approach is always available as its operational significance does not depend on the choices of approximators.

We will give below a metric on the space of probability measures on a common outcome space, which represents such an operational error measure. From the discussion above it is not surprising that in the commutative case the measure $\varepsilon_{\rm NO}$ is an over-estimate of this metric distance.

The trivial observables appearing in examples 2 and 3 do commute with the relevant target observables; but the measures $\varepsilon_{\rm NO}, \eta_{\rm NO}$ do then contain contributions from the preparation uncertainties; in more general commutative cases, they also include contributions arising from the specific underlying joint probability of A, C (in the case of $\varepsilon_{\rm NO}$). Thus, $\varepsilon_{\rm NO}, \eta_{\rm NO}$ depend on aspects of the specific chosen method of their determination and are not "pure" error or disturbance measures.

2.3 Unbiased approximations

If the measurement $\mathcal{M}$ is *unbiased*, that is, all first moments of C coincide with those of E^A, so that $C[x] = A$, then equation (6) reduces to

$$\varepsilon_{\text{NO}}(A, C, \rho)^2 = \langle C[x^2]\rangle_\rho - \langle C[x]^2\rangle_\rho = \langle C[x^2]\rangle_\rho - \langle A^2\rangle_\rho.$$

The quantity $V(C) := \langle C[x^2]\rangle_\rho - \langle C[x]^2\rangle_\rho$ indicates how much the POVM C deviates from being a projection valued (sharp) observable; we may refer to this quantity as the *intrinsic noise* of C. For two POVMs C, D that are jointly measurable, which means that they are marginals of a POVM G on the Cartesian product of the outcome sets, it is known that their intrinsic noise quantities satisfy an uncertainty relation [16–19]:

$$\Big(\langle C[x^2]\rangle_\rho - \langle C[x]^2\rangle_\rho\Big)\Big(\langle D[x^2]\rangle_\rho - \langle D[x]^2\rangle_\rho\Big) \geq \tfrac{1}{2}\Big|\langle [C[x], D[x]\rangle_\rho\Big|. \tag{8}$$

If the joint measurement G is taken as a simultaneous approximation of observables A and B and the condition of unbiasedness is satisfied for both observables, so that $A = C[1]$ and $B = D[1]$, this inequality can indeed be rewritten as

$$\varepsilon_{\text{NO}}(A, \mathcal{M}, \rho)\, \varepsilon_{\text{NO}}(B, \mathcal{M}, \rho) \;\geq\; \tfrac{1}{2}\Big|\langle [A, B\rangle_\rho\Big|. \tag{9}$$

However, the interpretation is not in terms of measurement errors in the first place: the inequality (8) rather states that if the jointly measurable observables C, D are unbiased approximators to the noncommuting observables A, B, respectively, then the degrees of sharpness of C and D are bounded by the commutator of A, B. In other words: *if the approximations are required to be good, in the sense of the absence of any systematic errors, then the approximating observables must be sufficiently unsharp as they are jointly measurable but noncommuting.*

Since now $C[x]^2 = A^2$ and $D[x]^2 = B^2$, the unsharpness measures on the left hand side of (8) do, *a fortiori*, provide operational estimates of differences between the target distributions, E^A_ρ and E^B_ρ, and the respective distributions of the measured observables C_ρ and D_ρ, namely, in terms of the second moments (where the first moments of E^A and C and those of E^B and D already coincide). It goes without saying that an assessment of the similarity of two probability distributions solely in terms of their second moments cannot be very accurate.

It is curious that one has to give up a feature that makes a measurement a good approximation – absence of systematic errors – in order to obtain a violation of (9). This observation in itself would appear to raise doubts over the general validity of the interpretation of ε_{NO} as a measure of error.

At this point we note that the inequality (3) can be understood as a special case of (9). In fact, the disturbance of B in an approximate measurement of A is an instance of an approximation error for B in terms of the POVM defined by making some measurement after the A measurement in order to obtain information about B. Such a sequence of measurements is equivalent to a joint measurement, so that the disturbance becomes itself an approximation error.

2.4 The three-state method: a resurrection of ε_{NO}?

In response to criticisms presented in [20, 21] that $\varepsilon_{\text{NO}}, \eta_{\text{NO}}$ generally lack operational significance, Ozawa [5] proposed a method of measuring ε_{NO} that was later termed *three-state method* [8]; it is encapsulated in the formula

$$\begin{aligned}\varepsilon_{\text{NO}}(A, \mathcal{M}, \rho)^2 &= \operatorname{tr}\rho A^2 + \operatorname{tr}\rho C[x^2] + \operatorname{tr}\rho C[x] + \operatorname{tr}\rho_1 C[x] - \operatorname{tr}\rho_2 C[x],\\ \rho_1 &= A\rho A, \quad \rho_2 = (A + \mathbf{1})\rho(A + \mathbf{1}),\end{aligned} \tag{10}$$

where C is the measured approximator observable. While now the quantity $\varepsilon_{\rm NO}$ is determined by the statistics of A and C, one can no longer claim that it is operationally significant with respect to a single state, ρ, thus undermining one of the perceived virtues of (4).

2.5 State-dependent measurement uncertainty relations – what use are they anyway?

This was the question my PhD student Neil Stevens asked upon reading our critical review of the noise operator approach [10], and he accompanied it with the following simple observation. Let A, B be two sharp observables, then for every state ρ_0 one can take trivial observables C, D defined as $C_\rho = A_{\rho_0}$ and $D_\rho = B_{\rho_0}$ for all ρ. Then for $\rho = \rho_0$, one obtains $C_\rho = A_\rho$ and $D_\rho = B_\rho$, that is, the given trivial observables provide the exact statistics of both A and B in just this state. Any faithful error measure should reflect this by assuming value 0 in that state.

This consideration teaches us two things. Firstly, one should not expect $|\langle[A, B]\rangle|$ to provide a nontrivial trade-off bound for state-dependent error measures $\Delta(A_\rho, C_\rho)$, $\Delta(B_\rho, D_\rho)$ as for every state there are the above trivial observables which make both such errors equal to zero, independently of whether the commutator expectation vanished.

Secondly, for these trivial observables eq. (6) gives $\varepsilon_{\rm NO}(A, \mathcal{M}, \rho_0) = \sqrt{2}\Delta(A, \rho_0)$. This should be an indication of the value-comparison error, but we see that ε comprises independent contributions from the preparation uncertainties not only of the probe (as the source of noise) but also from the preparation of the system itself. Hence, $\varepsilon_{\rm NO}$ fails to provide a “pure" error measure.

3 Quantum rms error

The original motivation for the noise operator based “error” may have been the observation that for an eigenstate ρ_a associated with an eigenvalue a of A it assumes the form of the classic Gaussian root-mean-square error:

$$\varepsilon_{\rm NO}(A, \mathcal{M}, \rho_a)^2 = \int (x-a)^2 C_{\rho_a}(dx),$$

where C is the observable measured by $\mathcal{M}$ that is taken as an approximator for A. We have seen how this formula generalises to the case where C commutes with A, so that the measure appearing in (7) is a proper probability measure; but we have also seen that this value-comparison error measure fails to be applicable in general. We now show that there is an alternative way of extending the notion of rms error to the quantum domain, one that meets the requirement of operational significance in full generality. This will be a distribution-comparison error measure that compares the relevant distributions.

3.1 Wasserstein 2-distance

The following defines a metric on the space of probability measures on $\mathbb{R}$: for two probability measures μ, ν we put

$$\Delta(\mu, \nu) = \left[\inf_\gamma \iint (x-x')^2 d\gamma(x, x')\right]^{\frac{1}{2}},$$

where the infimum is taken over all *couplings* γ of μ, ν (i.e., all probability measures on $\mathbb{R}^2$ with marginals μ, ν). This is the so-called Wasserstein distance of order 2.[2] It can be applied to the prob-

[2]More generally, the Wasserstein distance of order α is defined by replacing the exponents 2 and $1/2$ in the definition of $\mathcal{D}(\mu, \nu)$ by α and $1/\alpha$, respectively. It has been shown in [12] that in the case of position and momentum, joint measurement uncertainty relations follow for all choices of these deviations, with $\alpha \in [1, \infty]$.

ability measures A_ρ, C_ρ of two observables, defined as POVMs A, C on $\mathbb{R}$. By maximising the state-dependent error $\Delta(A_\rho, C_\rho)$ over all states ρ one has the worst-case error estimate,

$$\Delta(A, C) = \sup_\rho \Delta(A_\rho, C_\rho). \tag{11}$$

This is a metric on the space of observables on $\mathbb{R}$. It is evident that the measure is operationally significant since the inputs required for its calculation are the probability measures of the observables to be compared. The error $\Delta(A, C)$ is thus a figure of merit for all possible devices measuring an observable C as an approximate measurement of A.

3.2 Joint measurement uncertainty relations

As shown in [11] and [12], the following measurement error relation holds for approximate joint measurements of position Q and momentum P in terms of observable C, D:

$$\Delta(Q, C)\,\Delta(P, D) \geq \frac{\hbar}{2}. \tag{12}$$

Further, for approximate joint measurements of ±1-valued qubit observables A, B it has been shown in [13] that the approximation errors obey an additive trade-off relation of the form

$$\Delta(A, C)^2 + \Delta(B, D)^2 \geq (\text{incompatibility of } A, B). \tag{13}$$

Optimal approximations, in the sense of left hand side term reaching the lower bound, do exist and require, in particular, that the approximators have the same values, ±1, as the target observables.

3.3 Comparisons

Considering the form (7) of ε_{NO}, it is seen immediately that in the commutative case the bimeasure becomes a probability measure defined by $G(dxdy) = \operatorname{tr} \rho E^A(dx)C(dy)$; and this is a particular coupling of E^A_ρ and C_ρ, so that one obtains:

$$\varepsilon_{\text{NO}}(A, \mathcal{M}, \rho) \geq \Delta(A_\rho, C_\rho).$$

This shows that ε_{NO} is an over-estimate of the metric state-dependent error from the perspective of distribution-comparison error estimates; as a value-comparison error measure it is natural that ε_{NO} exceeds Δ.

An irony observed in [13] is that the experiments [8, 9] confirming Ozawa's inequality (4) use approximations for which the error $\varepsilon_{\text{NO}}(A, \mathcal{M}, \rho)$ and disturbance $\eta_{\text{NO}}(B, \mathcal{M}, \rho)$ are state-independent and in fact directly coupled to $\Delta(A, C)$ and $\Delta(B, D)$. As noted above, in such joint measurements the quantities ε_{NO} cease to be operationally significant and accessible. There are constellations of observables and states where these quantities are nonzero while the distributions in an individual state are identical, so that, say, $\Delta(A_\rho, C_\rho) = 0$. This confirms that ε_{NO}, η_{NO} do fail to indicate the absence of imprecision in an individual state that would be appropriate to distribution-comparison error measures.

In the case of qubit observables of the form measured in the Toronto experiment one has $\varepsilon_{\text{NO}}(A, \mathcal{M}, \rho) = \Delta(A, C)$ and $\eta_{\text{NO}}(B, \mathcal{M}, \rho) = \Delta(B, D)$. (In this case the target observables commute with their approximators.) Hence, we find again that the state dependence of the noise operator based measures has become obsolete; they have turned into perfect estimates of the worst-case metric errors. Moreover, as a consequence of (13), the ε_{NO} quantities now obey themselves a trade-off of the form (13), which refutes the alleged violation of a Heisenberg-type error-disturbance trade-off suggested by Ozawa's inequality (4).

4 Conclusion

We have reviewed two proposed error measures as candidates of a generalisation of the classical root-mean-square error adapted to quantum measurements. The first, propagated prominently by Ozawa, is naturally defined as an measure of average error of comparisons between the values obtained in a simultaneous measurement of the target observable and an approximating observable thereof. It was noted that this concept is only applicable in cases where the two observables in question commute. In the case of unbiased approximations this measure is found to become a fair estimate of the second error measure, defined via the Wasserstein 2-distance between probability distributions. This latter quantity is is always operationally significant and accessible, and its supremum over all states renders error measures as device figures of merit. Finally, we have given examples of Heisenberg-type joint measurement error trade-off relations in terms of these measures.

References

[1] W. Heisenberg, Zeitschr. Phys. **43**, 172 (1927)
[2] M. Ozawa, Phys. Lett. A **299**, 1 (2002)
[3] M. Ozawa, Phys. Lett. A **318**, 21 (2003)
[4] M. Ozawa, Phys. Rev. A **67**, 042105 (2003)
[5] M. Ozawa, Ann. Phys. **311**, 350 (2004)
[6] M. Ozawa, Phys. Lett. A **320**, 367 (2004)
[7] M. Ozawa, J. Opt. B **7**, S672 (2005)
[8] J. Erhart *et al*, Nature Phys. **8**, 185 (2012)
[9] L. Rozema *et al*, Phys. Rev. Lett. **109**, 100404 (2012)
[10] P. Busch, P. Lahti, R. Werner, arXiv:1312.4393 (2013)
[11] P. Busch, P. Lahti, R.F. Werner, Phys. Rev. Lett. **111**, 160405 (2013)
[12] P. Busch, P. Lahti, R.F. Werner, J. Math. Phys. **55**, 042111 (2014)
[13] P. Busch, P. Lahti, R.F. Werner, Phys. Rev. A **89**, 012129 (2014)
[14] W. Heisenberg, U. Chicago Press, Chicago p. 21 (1930)
[15] E. Kennard, Zeitschr. Phys. **44**, 326 (1927)
[16] E. Arthurs, J. Kelly, Bell Syst. Tech. J. **44**, 725 (1965)
[17] E. Arthurs, M. Goodman, Phys. Rev. Lett. **60**, 2447 (1988)
[18] S. Ishikawa, Rep. Math. Phys. **29**, 257 (1991)
[19] M. Ozawa, in *Quantum Aspects of Optical Communications*, eds. C. Bendjaballah, O. Hirota, S. Reynaud (Springer, Berlin Heidelberg, 1991), pp. 1–17
[20] P. Busch, T. Heinonen, P. Lahti, Phys. Lett. A **320**, 261 (2004)
[21] R.F. Werner, Quant. Inform. Comput. **4**, 546 (2004)