

# Uncertainty relations: An operational approach to the error-disturbance tradeoff

Joseph M. Renes<sup>1</sup>, Volkher B. Scholz<sup>1,2</sup>, and Stefan Huber<sup>1,3</sup>

<sup>1</sup>*Institute for Theoretical Physics, ETH Zürich, Switzerland*

<sup>2</sup>*Department of Physics, Ghent University, Belgium*

<sup>3</sup>*Department of Mathematics, Technische Universität München, Germany*

Accepted in *Quantum* July 10, 2017

The notions of error and disturbance appearing in quantum uncertainty relations are often quantified by the discrepancy of a physical quantity from its ideal value. However, these real and ideal values are not the outcomes of simultaneous measurements, and comparing the values of unmeasured observables is not necessarily meaningful according to quantum theory. To overcome these conceptual difficulties, we take a different approach and define error and disturbance in an operational manner. In particular, we formulate both in terms of the probability that one can successfully distinguish the actual measurement device from the relevant hypothetical ideal by any experimental test whatsoever. This definition itself does not rely on the formalism of quantum theory, avoiding many of the conceptual difficulties of usual definitions. We then derive new Heisenberg-type uncertainty relations for both joint measurability and the error-disturbance tradeoff for arbitrary observables of finite-dimensional systems, as well as for the case of position and momentum. Our relations may be directly applied in information processing settings, for example to infer that devices which can faithfully transmit information regarding one observable do not leak any information about conjugate observables to the environment. We also show that Englert's wave-particle duality relation [*Phys. Rev. Lett.* **77**, 2154 (1996)] can be viewed as an error-disturbance uncertainty relation.

## 1 Introduction

It is no overstatement to say that the uncertainty principle is a cornerstone of our understanding of quantum mechanics, clearly marking the departure of quantum physics from the world of classical physics. Heisenberg's original formulation in 1927 mentions two facets to the principle. The first restricts the joint measurability of observables, stating that noncommuting observables such as position and momentum can only be simultaneously determined with a characteristic amount of indeterminacy [1, p. 172] (see [2, p. 62] for an English translation). The second describes an error-disturbance tradeoff, noting that the more precise a measurement of one observable is made, the greater the disturbance to noncommuting observables [1, p. 175] ([2, p. 64]). The two are of course closely related, and Heisenberg argues for the former on the basis of the latter. Neither version can be taken merely as a limitation on measurement of otherwise well-defined values of position and momentum, but rather as questioning the sense in which values of two noncommuting observables can even be said to simultaneously exist. Unlike classical mechanics, in the framework of quantum mechanics we cannot necessarily regard unmeasured quantities as physically meaningful.

More formal statements were constructed only much later, due to the lack of a precise mathematical description of the measurement process in quantum mechanics. Here we must be careful to draw a distinction between statements addressing Heisenberg's original notions of uncertainty from those, like the standard Kennard-Robertson uncertainty relation [3, 4], which address the impossibility of finding a quantum state with well-defined values for noncommuting observables. Entropic uncertainty relations [5, 6] are also an example of this class; see [7] for a review. Joint measurability has a longer history, going back at least to the seminal work of Arthurs and Kelly [8] and continuing in [9–27]. Quantitative error-disturbance relations have only been formulated relatively recently, going back at least to Braginsky and Khalili [28, Chap. 5] and continuing in [20, 29–35].

Beyond technical difficulties in formulating uncertainty relations, there is a perhaps more difficult conceptual hurdle in that the intended consequences of the uncertainty principle seem to preclude their own straightforward formalization. To find a relation between, say, the error of a position measurement and its disturbance to momentum in a given experimental setup like the gamma ray microscope would seem to require comparing the actual values of position and momentum with their supposed ideal values. However, according to the uncertainty principle itself, we should be wary of simultaneously ascribing well-defined values to the actual and ideal position and momentum since they do not correspond to commuting observables. Thus, it is not immediately clear how to formulate either meaningful measures of error and disturbance, for instance as mean-square deviations between real and ideal values, or a meaningful relation between them.<sup>1</sup> This question is the subject of much ongoing debate [25, 30, 36–39].

<sup>1</sup>Uncertainty relations like the Kennard-Robertson bound or entropic relations do not face this issue as they do not attempt to compare actual and ideal values of the observables.

Without drawing any conclusions as to the ultimate success or failure of this program, in this paper we propose a completely different approach which we hope sheds new light on these conceptual difficulties. Here, we define error and disturbance in an operational manner and ask for uncertainty relations that are statements about the properties of measurement devices, not of fixed experimental setups or of physical quantities themselves. More specifically, we define error and disturbance in terms of the *distinguishing probability*, the probability that the actual behavior of the measurement apparatus can be distinguished from the relevant ideal behavior in any single experiment whatsoever. To characterize measurement error, for example, we imagine a black box containing either the actual device or the ideal device. By controlling the input and observing the output we can make an informed guess as to which is the case. We then attribute a large measurement error to the measurement apparatus if it is easy to tell the difference, so that there is a high probability of correctly guessing, and a low error if not; of course we pick the optimal input states and output measurements for this purpose. In this way we do not need to attribute a particular ideal value of the observable to be measured, we do not need to compare actual and ideal values themselves (nor do we necessarily even care what the possible values are), and instead we focus squarely on the properties of the device itself. Intuitively, we might expect that calibration provides the strictest test, i.e. inputting states with a known value of the observable in question. But in fact this is not the case, as entanglement at the input can increase the distinguishability of two measurements. The merit of this approach is that the notion of distinguishability itself does not rely on any concepts or formalism of quantum theory, which helps avoid conceptual difficulties in formalizing the uncertainty principle.

Defining the disturbance an apparatus causes to an observable is more delicate, as an observable itself does not have a directly operational meaning (as opposed to the measurement of an observable). But we can consider the disturbance made either to an ideal measurement of the observable or to ideal preparation of states with well-defined values of the observable. In all cases, the error and disturbance measures we consider are directly linked to a well-studied norm on quantum channels known as the completely bounded norm or diamond norm. We can then ask for bounds on the error and disturbance quantities for two given observables that every measurement apparatus must satisfy. In particular, we are interested in bounds depending only on the chosen observables and not the particular device. Any such relation is a statement about measurement devices themselves and is not specific to the particular experimental setup in which they are used. Nor are such relations statements about the values or behavior of physical quantities themselves. In this sense, we seek statements of the uncertainty principle akin to Kelvin's form of the second law of thermodynamics as a constraint on thermal machines, and not like Clausius's or Planck's form involving the behavior of physical quantities (heat and entropy, respectively). By appealing to a fundamental constraint on quantum dynamics, the continuity (in the completely bounded norm) of the Stinespring dilation [40, 41], we find error-disturbance uncertainty relations for arbitrary observables in finite dimensions, as well as for position and momentum. Furthermore, we show how the relation for measurement error and measurement disturbance can be transformed into a joint-measurability uncertainty relation. Interestingly, we also find that Englert's wave-particle duality relation [42] can be viewed as an error-disturbance relation.

The case of position and momentum illustrates the stark difference between the kind of uncertainty statements we can make in our approach with one based on the notion of comparing real and ideal values. Take the notion of joint measurability, where we would like to formalize the notion that no device can accurately measure both position and momentum. In the latter approach one would first try to quantify the amount of position or momentum error made by a device as the discrepancy to the true value, and then show that they cannot both be small. The errors would be in units of position or momentum, respectively, and the hoped-for uncertainty relation would pertain to these values. Here, in contrast, we focus on the performance of the actual device relative to fixed ideal devices, in this case idealized separate measurements of position or momentum. Importantly, we need not think of the ideal measurement as having infinite precision. Instead, we can pick any desired precision and ask if the behavior of the actual device is essentially the same as this precision-limited ideal. Now the position and momentum errors do not have units of these quantities (they are unitless and always lie between zero and one), but instead *depend on the desired precision*. Our uncertainty relation then implies that both errors cannot be small if we demand high precision in both position and momentum. In particular, when the product of the scales of the two precisions is small compared to Planck's constant, then the errors will be bounded away from zero (see Theorem 3 for a precise statement). It is certainly easier to have a small error in this sense when the demanded precision is low, and this accords nicely with the fact that sufficiently-inaccurate joint measurement is possible. Indeed, we find no bound on the errors for low precision.

An advantage and indeed a separate motivation of an operational approach is that bounds involving operational quantities are often useful in analyzing information processing protocols. For example, entropic uncertainty relations, which like the Robertson relation characterize quantum states, have proven very useful

in establishing simple proofs of the security of quantum key distribution [6, 7, 43–45]. Here we show that the error-disturbance relation implies that quantum channels which can faithfully transmit information regarding one observable do not leak any information whatsoever about conjugate observables to the environment. This statement cannot be derived from entropic relations, as it holds for all channel inputs. It can be used to construct leakage-resilient classical computers from fault-tolerant quantum computers [46], for instance.

The remainder of the paper is structured as follows. In the next section we give the mathematical background necessary to state our results, and describe how the general notion of distinguishability is related to the completely bounded norm (cb norm) in this setting. In Section 3 we define our error and disturbance measures precisely. Section 4 presents the error-disturbance tradeoff relations for finite dimensions, and details how joint measurability relations can be obtained from them. Section 5 considers the error-disturbance tradeoff relations for position and momentum. Two applications of the tradeoffs are given in Section 6: a formal statement of the information disturbance tradeoff for information about noncommuting observables and the connection between error-disturbance tradeoffs and Englert’s wave-particle duality relations. In Section 7 we compare our results to previous approaches in more detail, and finally we finish with open questions in Section 8.

## 2 Mathematical setup

### 2.1 Distinguishability

The notion of the distinguishing probability is independent of the mathematical framework needed to describe quantum systems, so we give it first. Consider an apparatus  $\mathcal{E}$  which in some way transforms an input  $A$  into an output  $B$ . To describe how different  $\mathcal{E}$  is from another such apparatus  $\mathcal{E}'$ , we can imagine the following scenario. Suppose that we randomly place either  $\mathcal{E}$  or  $\mathcal{E}'$  into a black box such that we no longer have any access to the inner workings of the device, only its inputs and outputs. Now our task is to guess which device is actually in the box by performing a single experiment, feeding in any desired input and observing the output in any manner of our choosing. In particular, the inputs and measurements can and should depend on  $\mathcal{E}$  and  $\mathcal{E}'$ . The probability of making a correct guess, call it  $p_{\text{dist}}(\mathcal{E}, \mathcal{E}')$ , ranges from  $\frac{1}{2}$  to 1, since we can always just make a random guess without doing any experiment on the box at all. Therefore it is more convenient to work with the distinguishability measure

$$\delta(\mathcal{E}, \mathcal{E}') := 2p_{\text{dist}}(\mathcal{E}, \mathcal{E}') - 1, \quad (1)$$

which ranges from zero (completely indistinguishable) to one (completely distinguishable). Later on we will show this quantity takes a specific mathematical form in quantum mechanics. But note that the definition implies that the distinguishability is monotonic under concatenation with a channel  $\mathcal{F}$  to both  $\mathcal{E}$  and  $\mathcal{E}'$ , since this just restricts the possible tests. That is, both  $\delta(\mathcal{E}\mathcal{F}, \mathcal{E}'\mathcal{F}) \leq \delta(\mathcal{E}, \mathcal{E}')$  and  $\delta(\mathcal{F}\mathcal{E}, \mathcal{F}\mathcal{E}') \leq \delta(\mathcal{E}, \mathcal{E}')$  hold for all channels  $\mathcal{F}$  whose inputs and outputs are such that the channel concatenation is sensible. Here and in the remainder of the paper, we denote concatenation of channels by juxtaposition, while juxtaposition of operators denotes multiplication as usual.

### 2.2 Systems, algebras, channels, and measurements

In the finite-dimensional case we will be interested in two arbitrary nondegenerate observables denoted  $X$  and  $Z$ . Only the eigenvectors of the observables will be relevant, call them  $|\varphi_x\rangle$  and  $|\theta_z\rangle$ , respectively. In infinite dimensions we will confine our analysis to position  $Q$  and momentum  $P$ , taking  $\hbar = 1$ . The analog of  $Q$  and  $P$  in finite dimensions are canonically conjugate observables  $X$  and  $Z$  for which  $|\varphi_x\rangle = \frac{1}{\sqrt{d}} \sum_z \omega^{xz} |\theta_z\rangle$ , where  $d$  is the dimension and  $\omega$  is a primitive  $d$ th root of unity.

It will be more convenient for our purposes to adopt the algebraic framework and use the Heisenberg picture, though we shall occasionally employ the Schrödinger picture. In the Heisenberg picture we describe systems chiefly by the algebra of observables on them and describe transformations of systems by quantum channels, completely positive and unital maps from the algebra of observables of the output to the observables of the input [10, 47–50]. This allows us to treat classical and quantum systems on an equal footing within the same framework. When the input or output system is quantum mechanical, the observables are the bounded operators  $\mathcal{B}(\mathcal{H})$  from the Hilbert space  $\mathcal{H}$  associated with the system to itself. Classical systems, such as the results of measurement or inputs to a state preparation device, take values in a set, call it  $Y$ . The relevant algebra of observables here is  $L^\infty(Y)$ , the (bounded, measurable) functions on  $Y$ . Hybrid systems are described by tensor products, so an apparatus  $\mathcal{E}$  which measures a quantum system has an output algebra described by  $L^\infty(Y) \otimes \mathcal{B}(\mathcal{H})$ . To describe just the measurement result, we keep only  $L^\infty(Y)$ . We shall occasionally denote the input and output spaces explicitly as  $\mathcal{E}_{A \rightarrow YB}$  when useful.

For arbitrary input and output algebras  $\mathcal{A}_A$  and  $\mathcal{A}_B$ , quantum channels are precisely those maps  $\mathcal{E}$  which are unital,  $\mathcal{E}(\mathbb{1}_B) = \mathbb{1}_A$ , and completely positive, meaning that not only does  $\mathcal{E}$  map positive elements of  $\mathcal{A}_B$  to positive elements of  $\mathcal{A}_A$ , it also maps positive elements of  $\mathcal{A}_B \otimes \mathcal{B}(\mathbb{C}^n)$  to positive elements of  $\mathcal{A}_A \otimes \mathcal{B}(\mathbb{C}^n)$  for all integer  $n$ . This requirement is necessary to ensure that channels act properly on entangled systems.

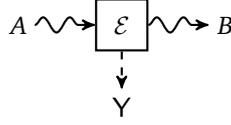


Figure 1: A general quantum apparatus  $\mathcal{E}$ . The apparatus measures a quantum system  $A$  giving the output  $Y$ . In so doing,  $\mathcal{E}$  also transforms the input  $A$  into the output system  $B$ . Here the wavy lines denote quantum systems, the dashed lines classical systems. Formally, the apparatus is described by a quantum instrument.

A general measurement apparatus has both classical and quantum outputs, corresponding to the measurement result and the post-measurement quantum system. Channels describing such devices are called *quantum instruments*; we will call the channel describing just the measurement outcome a *measurement*. In finite dimensions any measurement can be seen as part of a quantum instrument, but not so for idealized position or momentum measurements, as shown in Theorem 3.3 of [10] (see page 57). Technically, we may anticipate the result since the post-measurement state of such a device would presumably be a delta function located at the value of the measurement, which is not an element of  $L^2(Q)$ . This need not bother us, though, since it is not operationally meaningful to consider a position measurement instrument of infinite precision. And indeed there is no mathematical obstacle to describing finite-precision position measurement by quantum instruments, as shown in Theorem 6.1 (page 67 of [10]). For any bounded function  $\alpha \in L^2(Q)$  we can define the instrument  $\mathcal{E}_\alpha : L^\infty(Q) \otimes \mathcal{B}(\mathcal{H}) \rightarrow \mathcal{B}(\mathcal{H})$  by

$$\mathcal{E}_\alpha(f \otimes a) = \int dq f(q) A_{q;\alpha}^* a A_{q;\alpha}, \quad (2)$$

where  $A_{q;\alpha}\psi(q') = \alpha(q - q')\psi(q')$  for all  $\psi \in L^2(Q)$ . The classical output of the instrument is essentially the ideal value convolved with the function  $\alpha$ . Thus, setting the width of  $\alpha$  sets the precision limit of the instrument.

### 2.3 Distinguishability as a channel norm

The distinguishability measure is actually a norm on quantum channels, equal (apart from a factor of one half) to the so-called norm of complete boundedness, the cb norm [51–53]. The cb norm is defined as an extension of the operator norm, similar to the extension of positivity above, as

$$\|T\|_{\text{cb}} := \sup_{n \in \mathbb{N}} \|\mathbb{1}_n \otimes T\|_\infty, \quad (3)$$

where  $\|T\|_\infty$  is the operator norm. Then

$$\delta(\mathcal{E}_1, \mathcal{E}_2) = \frac{1}{2} \|\mathcal{E}_1 - \mathcal{E}_2\|_{\text{cb}}. \quad (4)$$

In the Schrödinger picture we instead extend the trace norm  $\|\cdot\|_1$ , and the result is usually called the diamond norm [51, 53]. In either case, the extension serves to account for entangled inputs in the experiment to test whether  $\mathcal{E}_1$  or  $\mathcal{E}_2$  is the actual channel. In fact, entanglement is helpful even when the channels describe projective measurements, as shown by an example given in Appendix A. This expression for the cb or diamond norm is not closed-form, as it requires an optimization. However, in finite dimensions the cb norm can be cast as a convex optimization, specifically as a semidefinite program [54, 55], which makes numerical computation tractable. Further details are given in Appendix B.

### 2.4 The Stinespring representation and its continuity

According to the Stinespring representation theorem [52, 56], any channel  $\mathcal{E}$  mapping an algebra  $\mathcal{A}$  to  $\mathcal{B}(\mathcal{H})$  can be expressed in terms of an isometry  $V : \mathcal{H} \rightarrow \mathcal{K}$  to some Hilbert space  $\mathcal{K}$  and a representation  $\pi$  of  $\mathcal{A}$  in  $\mathcal{B}(\mathcal{K})$  such that, for all  $a \in \mathcal{A}$ ,

$$\mathcal{E}(a) = V^* \pi(a) V. \quad (5)$$

The isometry in the Stinespring representation is usually called the *dilation* of the channel, and  $\mathcal{K}$  the dilation space. In finite-dimensional settings, calling the input  $A$  and the output  $B$ , one usually considers maps taking  $\mathcal{A} = \mathcal{B}(\mathcal{H}_B)$  to  $\mathcal{B}(\mathcal{H}_A)$ . Then one can choose  $\mathcal{K} = \mathcal{H}_B \otimes \mathcal{H}_E$ , where  $\mathcal{H}_E$  is a suitably large Hilbert space associated to the “environment” of the transformation ( $\mathcal{H}_E$  can always be chosen to have dimension  $\dim(\mathcal{H}_A) \dim(\mathcal{H}_B)$ ). The representation  $\pi$  is just  $\pi(a) = a \otimes \mathbb{1}_E$ . Using the isometry  $V$ , we can also construct a channel from  $\mathcal{B}(\mathcal{H}_E)$  to  $\mathcal{B}(\mathcal{H}_A)$  in the same manner; this is known as the complement  $\mathcal{E}^\#$  of  $\mathcal{E}$ .

The advantage of the general form of the Stinespring representation is that we can easily describe measurements, possibly continuous-valued, as well. For the case of finite outcomes, consider the ideal projective measurement  $\mathcal{Q}_X$  of the observable  $X$ . Choosing a basis  $\{|b_x\rangle\}$  of  $L^2(X)$  and defining  $\pi(\delta_x) = |b_x\rangle\langle b_x|$  for  $\delta_x$  the function taking the value 1 at  $x$  and zero elsewhere, the canonical dilation isometry  $W_X : \mathcal{H} \rightarrow L^2(X) \otimes \mathcal{H}$  is given by

$$W_X = \sum_x |b_x\rangle \otimes |\varphi_x\rangle\langle\varphi_x|. \quad (6)$$

Note that this isometry defines a quantum instrument, since it can describe both the measurement outcome and the post-measurement quantum system. If we want to describe just the measurement result, we could simply use  $W_X = \sum_x |b_x\rangle \langle\varphi_x|$  with the same  $\pi$ . More generally, a POVM with elements  $\Lambda_x$  has the isometry  $W_X = \sum_x |b_x\rangle \langle\varphi_x| \sqrt{\Lambda_x}$ .

For finite-precision measurements of position or momentum, the form of the quantum instrument in (2) immediately gives a Stinespring dilation  $W_Q : \mathcal{H} \rightarrow \mathcal{K}$  with  $\mathcal{K} = L^2(Q) \otimes \mathcal{H}$  whose action is defined by

$$(W_Q\psi)(q, q') = \alpha(q - q')\psi(q'), \quad (7)$$

and where  $\pi$  is just pointwise multiplication on the  $L^\infty(Q)$  factor, i.e. for  $f \in L^\infty(Q)$ , and  $a \in \mathcal{B}(\mathcal{H})$ ,  $[\pi(f \otimes a)(\xi \otimes \psi)](q, q') = f(q)\xi(q) \cdot (a\psi)(q')$  for all  $\xi \in L^2(Q)$  and  $\psi \in \mathcal{H}$ .

A slight change to the isometry in (6) gives the dilation of the device which prepares the state  $|\varphi_x\rangle$  for classical input  $x$ . Formally the device is described by the map  $\mathcal{P} : \mathcal{B}(\mathcal{H}) \rightarrow L^2(X)$  for which  $\mathcal{P}(\Lambda) = \sum_x |b_x\rangle\langle b_x| \langle\varphi_x| \Lambda |\varphi_x\rangle$ . Now consider  $W'_X : L^2(X) \rightarrow \mathcal{H} \otimes L^2(X)$  given by

$$W'_X = \sum_x |\varphi_x\rangle \otimes |b_x\rangle\langle b_x|. \quad (8)$$

Choosing  $\pi(\Lambda) = \Lambda \otimes \mathbb{1}_X$ , we have  $\mathcal{P}(\Lambda) = W_X'^* \pi(\Lambda) W_X'$ .

The Stinespring representation is not unique [41]. Given two representations  $(\pi_1, V_1, \mathcal{K}_1)$  and  $(\pi_2, V_2, \mathcal{K}_2)$  of the same channel  $\mathcal{E}$ , there exists a partial isometry  $U : \mathcal{K}_1 \rightarrow \mathcal{K}_2$  such that  $UV_1 = V_2$ ,  $U^*V_2 = V_1$ , and  $U\pi_1(a) = \pi_2(a)U$  for all  $a \in \mathcal{A}$ . For the representations  $\pi$  as usually employed for the finite-dimensional case, this last condition implies that  $U$  is a partial isometry from one environment to the other, for  $U(a \otimes \mathbb{1}_E) = (a \otimes \mathbb{1}_{E'})U$  can only hold for all  $a$  if  $U$  acts trivially on  $B$ . For channels describing measurements, finite or continuous, the last condition implies that any such  $U$  is a conditional partial isometry, dependent on the outcome of the measurement result. Thus, for any set of isometries  $U_x : \mathcal{H}_S \rightarrow \mathcal{H}_R$ ,  $\sum_x |b_x\rangle \otimes U_x |\varphi_x\rangle\langle\varphi_x| U_x^*$  is a valid dilation of  $\mathcal{Q}_X$ , just as is  $W_X$  in (6). Similarly,  $(W_Q'\psi)(q, q') = \alpha(q - q')[U_q\psi](q')$  is a valid dilation of  $\mathcal{E}_\alpha$  in (2).

The main technical ingredient required for our results is the continuity of the Stinespring representation in the cb norm [40, 41]. That is, channels which are nearly indistinguishable have Stinespring dilations which are close and vice versa. For completely positive and unital maps  $\mathcal{E}_1$  and  $\mathcal{E}_2$ , [40, 41] show that

$$\frac{1}{2} \|\mathcal{E}_1 - \mathcal{E}_2\|_{\text{cb}} \leq \inf_{\pi_i, V_i} \|V_1 - V_2\|_\infty \leq \sqrt{\|\mathcal{E}_1 - \mathcal{E}_2\|_{\text{cb}}}, \quad (9)$$

where the infimum is taken over all Stinespring representations  $(\pi_i, V_i, \mathcal{K}_i)$  of  $\mathcal{E}_i$ .

## 2.5 Sequential and joint measurements

Using the Stinespring representation we can easily show that, in principle, any joint measurement can always be decomposed into sequential measurement.

**Lemma 1.** *Suppose that  $\mathcal{E} : L^\infty(X) \otimes L^\infty(Z) \rightarrow \mathcal{B}(\mathcal{H})$  is a channel describing a joint measurement. Then there exists an apparatus  $\mathcal{A} : L^\infty(X) \otimes \mathcal{B}(\mathcal{H}') \rightarrow \mathcal{B}(\mathcal{H})$  and a conditional measurement  $\mathcal{M} : L^\infty(X) \otimes L^\infty(Z) \rightarrow L^\infty(X) \otimes \mathcal{B}(\mathcal{H}')$  such that  $\mathcal{E} = \mathcal{A}\mathcal{M}$ .*

*Proof.* Define  $\mathcal{M}' : L^\infty(X) \rightarrow \mathcal{B}(\mathcal{H})$  to be just the  $X$  output of  $\mathcal{E}$ , i.e.  $\mathcal{M}'(f) = \mathcal{E}(f \otimes 1)$ . Now suppose that  $V : \mathcal{H} \rightarrow L^2(X) \otimes L^2(Z) \otimes \mathcal{H}'$  is a Stinespring representation of  $\mathcal{E}$  and  $V_X : \mathcal{H} \rightarrow L^2(X) \otimes \mathcal{H}'$  is a representation of  $\mathcal{M}'$ , both with the standard representation  $\pi$  of  $L^\infty$  into  $L^2$ . By construction,  $V$  is also a dilation of  $\mathcal{M}'$ , and therefore there exists a partial isometry  $U_X$  such that  $V = U_X V_X$ . More specifically, conditional on the value  $X = x$ , each  $U_x$  sends  $\mathcal{H}'$  to  $L^2(Z) \otimes \mathcal{H}''$ . Thus, setting  $\mathcal{A}(f \otimes a) = V_X^*(\pi(f) \otimes a)V_X$  and  $\mathcal{M}_x(f) = U_x^*(\pi(f) \otimes \mathbb{1})U_x$ , we have  $\mathcal{E} = \mathcal{A}\mathcal{M}$ .  $\square$

### 3 Definitions of error and disturbance

#### 3.1 Measurement error

To characterize the error  $\varepsilon_X$  an apparatus  $\mathcal{E}$  makes relative to an ideal measurement  $\mathcal{Q}_X$  of an observable  $X$ , we can simply use the distinguishability of the two channels, taking only the classical output of  $\mathcal{E}$ . Suppose that the apparatus is described by the channel  $\mathcal{E} : \mathcal{B}(\mathcal{H}_B) \otimes L^\infty(X) \rightarrow \mathcal{B}(\mathcal{H}_A)$  and the ideal measurement by the channel  $\mathcal{Q}_X : L^\infty(X) \rightarrow \mathcal{B}(\mathcal{H}_A)$ . To ignore the output system  $B$ , we make use of the partial trace map  $\mathcal{T}_B : L^\infty(X) \rightarrow \mathcal{B}(\mathcal{H}_B) \otimes L^\infty(X)$  given by  $\mathcal{T}_B(f) = \mathbb{1}_B \otimes f$ . Then a sensible notion of error is given by  $\varepsilon_X(\mathcal{E}) = \delta(\mathcal{Q}_X, \mathcal{E}\mathcal{T}_B)$ . If it is easy to tell the ideal measurement apart from the actual device, then the error is large; if it is difficult, then the error is small.

As a general definition, though, this quantity is deficient to two respects. First, we could imagine an apparatus which performs an ideal  $\mathcal{Q}_X$  measurement, but simply mislabels the outputs. This leads to  $\varepsilon_X(\mathcal{E}) = 1$ , even though the ideal measurement is actually performed. Second, we might wish to consider the case that the classical output set of the apparatus is not equal to  $X$  itself. For instance, perhaps  $\mathcal{E}$  delivers much more output than is expected from  $\mathcal{Q}_X$ . In this case we also formally have  $\varepsilon_X(\mathcal{E}) = 1$ , since we can just examine the output to distinguish the two devices.

We can remedy both of these issues by describing the apparatus by the channel  $\mathcal{E} : \mathcal{B}(\mathcal{H}_B) \otimes L^\infty(Y) \rightarrow \mathcal{B}(\mathcal{H}_A)$  and just including a further classical postprocessing operation  $\mathcal{R} : L^\infty(Y) \rightarrow L^\infty(X)$  in the distinguishability step. Since we are free to choose the best such map, we define

$$\varepsilon_X(\mathcal{E}) := \inf_{\mathcal{R}} \delta(\mathcal{Q}_X, \mathcal{E}\mathcal{R}\mathcal{T}_B). \quad (10)$$

The setup of the definition is depicted in Figure 2.

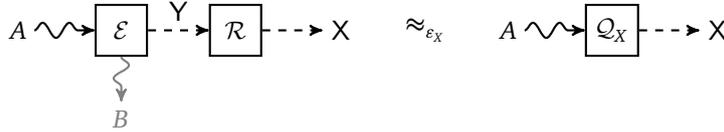


Figure 2: Measurement error. The error made by the apparatus  $\mathcal{E}$  in measuring  $X$  is defined by how distinguishable the actual device is from the ideal measurement  $\mathcal{Q}_X$  in any experiment whatsoever, after suitably processing the classical output  $Y$  of  $\mathcal{E}$  with the map  $\mathcal{R}$ . To enable a fair comparison, we ignore the quantum output of the apparatus, indicated in the diagram by graying out  $B$ . If the actual and ideal devices are difficult to tell apart, the error is small.

#### 3.2 Measurement disturbance

Defining the disturbance an apparatus  $\mathcal{E}$  causes to an observable, say  $Z$ , is more delicate, as an observable itself does not have a directly operational meaning. But there are two straightforward ways to proceed: we can either associate the observable with measurement or with state preparation. In the former, we compare how well we can mimic the ideal measurement  $\mathcal{Q}_Z$  of the observable after employing the apparatus  $\mathcal{E}$ , quantifying this using measurement error as before. Additionally, we should allow the use of recovery operations in which we attempt to “restore” the input state as well as possible, possibly conditional on the output of the measurement. Formally, let  $\mathcal{Q}_Z : L^\infty(Z) \rightarrow \mathcal{B}(\mathcal{H}_A)$  be the ideal  $Z$  measurement and  $\mathcal{R}$  be a recovery map  $\mathcal{R} : \mathcal{B}(\mathcal{H}_A) \rightarrow \mathcal{B}(\mathcal{H}_B) \otimes L^\infty(X)$  which acts on the output of  $\mathcal{E}$  conditional on the value of the classical output  $X$  (which it then promptly forgets). As depicted in Figure 3, the measurement disturbance is then the measurement error after using the best recovery map:

$$v_Z(\mathcal{E}) := \inf_{\mathcal{R}} \delta(\mathcal{Q}_Z, \mathcal{E}\mathcal{R}\mathcal{T}_Y\mathcal{Q}_Z). \quad (11)$$

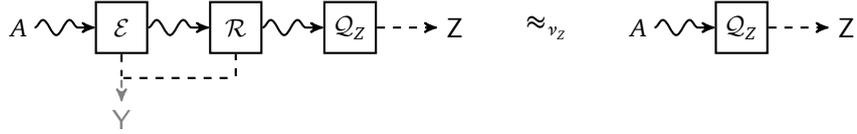


Figure 3: Measurement disturbance. To define the disturbance imparted by an apparatus  $\mathcal{E}$  to the measurement of an observable  $Z$ , consider performing the ideal  $\mathcal{Q}_Z$  measurement on the output  $B$  of  $\mathcal{E}$ . First, however, it may be advantageous to “correct” or “recover” the original input  $A$  by some operation  $\mathcal{R}$ . In general,  $\mathcal{R}$  may depend on the output  $X$  of  $\mathcal{E}$ . The distinguishability between the resulting combined operation and just performing  $\mathcal{Q}_Z$  on the original input defines the measurement disturbance.

### 3.3 Preparation disturbance

For state preparation, consider a device with classical input and quantum output that prepares the eigenstates of  $Z$ . We can model this by a channel  $\mathcal{P}_Z$ , which in the Schrödinger picture produces  $|\theta_z\rangle$  upon receiving the input  $z$ . Now we compare the action of  $\mathcal{P}_Z$  to the action of  $\mathcal{P}_Z$  followed by  $\mathcal{E}$ , again employing a recovery operation. Formally, let  $\mathcal{P}_Z : \mathcal{B}(\mathcal{H}_A) \rightarrow L^\infty(Z)$  be the ideal  $Z$  preparation device and consider recovery operations  $\mathcal{R}$  of the form  $\mathcal{R} : \mathcal{B}(\mathcal{H}_A) \rightarrow \mathcal{B}(\mathcal{H}_B) \otimes L^\infty(X)$ . Then the preparation disturbance is defined as

$$\eta_Z(\mathcal{E}) := \inf_{\mathcal{R}} \delta(\mathcal{P}_Z, \mathcal{P}_Z \mathcal{E} \mathcal{R} \mathcal{T}_Y). \quad (12)$$

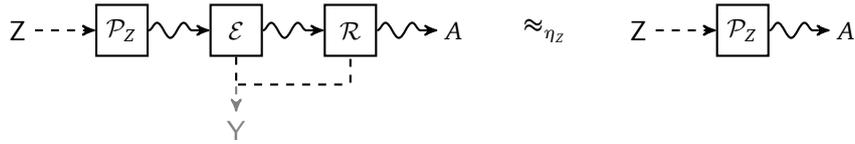


Figure 4: Preparation disturbance. The ideal preparation device  $\mathcal{P}_Z$  takes a classical input  $Z$  and creates the corresponding  $Z$  eigenstate. As with measurement disturbance, the preparation disturbance is related to the distinguishability of the ideal preparation device  $\mathcal{P}_Z$  and  $\mathcal{P}_Z$  followed by the apparatus  $\mathcal{E}$  in question and the best possible recovery operation  $\mathcal{R}$ .

All of the measures defined so far are “figures of merit”, in the sense that we compare the actual device to the ideal, perfect functionality. In the case of state preparation we can also define a disturbance measure as a “figure of demerit”, by comparing the actual functionality not to the best-case behavior but to the worst. To this end, consider a state preparation device  $\mathcal{C}$  which just ignores the classical input and always prepares the same fixed output state. These are constant (output) channels, and clearly  $\mathcal{E}$  disturbs the state preparation  $\mathcal{P}_Z$  considerably if  $\mathcal{P}_Z \mathcal{E}$  has effectively a constant output. Based on this intuition, we can then make the following formal definition:

$$\hat{\eta}_Z(\mathcal{E}) := \frac{d-1}{d} - \inf_{\mathcal{C}: \text{const.}} \delta(\mathcal{C}, \mathcal{P}_Z \mathcal{E}). \quad (13)$$

The disturbance is small according to this measure if it is easy to distinguish the action of  $\mathcal{P}_Z \mathcal{E}$  from having a constant output, and large otherwise. To see that  $\hat{\eta}_Z$  is positive, use the Schrödinger picture and let the output of  $\mathcal{C}^*$  be the state  $\sigma$  for all inputs. Then note that  $\inf_{\mathcal{C}} \delta(\mathcal{C}, \mathcal{P}_Z \mathcal{E}) = \min_{\mathcal{C}} \max_z \delta(\sigma, \mathcal{E}^*(\theta_z))$ , where the latter  $\delta$  is the trace distance. Choosing  $\sigma = \frac{1}{d} \sum_z \mathcal{E}^*(\theta_z)$  and using joint convexity of the trace distance, we have  $\inf_{\mathcal{C}} \delta(\mathcal{C}, \mathcal{P}_Z \mathcal{E}) \leq \frac{d-1}{d}$ .

We remark that while this disturbance measure leads to finite bounds in the case of finite dimensions, it is less well behaved in the case of position and momentum measurements: Without any bound on the energy of the test states, two channels tend to be as distinguishable as possible, unless they are already constant channels. To be more precise, any non-constant channel which only changes the energy by a fixed amount can be differentiated from a constant channel by inputting states of very high energy. Roughly speaking, even an arbitrarily strongly disturbing operation can be used to gain *some* information about the input and hence a constant channel is not a good “worst case” scenario. This is in sharp contrast to the finite-dimensional case, and supports the view that the disturbance measures  $\nu_Z(\mathcal{E})$  and  $\eta_Z(\mathcal{E})$  are physically more sensible.

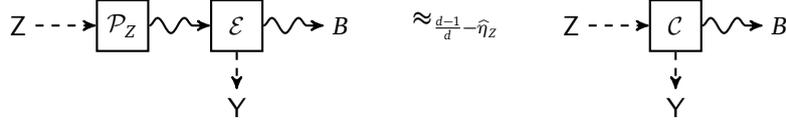


Figure 5: Figure of “demerit” version of preparation disturbance. Another approach to defining preparation disturbance is to consider distinguishability to a non-ideal device instead of an ideal device. The apparatus  $\mathcal{E}$  imparts a large disturbance to the preparation  $\mathcal{P}_Z$  if the output of the combination  $\mathcal{P}_Z\mathcal{E}$  is essentially independent of the input. Thus we consider the distinguishability of  $\mathcal{P}_Z\mathcal{E}$  and a constant preparation  $\mathcal{C}$  which outputs a fixed state regardless of the input  $Z$ .

For finite-dimensional systems, all the measures of error and disturbance can be expressed as semidefinite programs, as detailed in Appendix B. As an example, we compute these measures for the simple case of a non-ideal  $X$  measurement on a qubit; we will meet this example later in assessing the tightness of the uncertainty relations and their connection to wave-particle duality relations in the Mach-Zehnder interferometer. Consider the ideal measurement isometry (6), and suppose that the basis states  $|b_x\rangle$  are replaced by two pure states  $|\gamma_x\rangle$  which have an overlap  $\langle\gamma_0|\gamma_1\rangle = \sin\theta$ . Without loss of generality, we can take  $|\gamma_x\rangle = \cos\frac{\theta}{2}|b_x\rangle + \sin\frac{\theta}{2}|b_{x+1}\rangle$ . The optimal measurement  $\mathcal{Q}$  for distinguishing these two states is just projective measurement in the  $|b_x\rangle$  basis, so let us consider the channel  $\mathcal{E}_{\text{MZ}} = \mathcal{W}\mathcal{Q}$ . Then, as detailed in Appendix B, for  $Z$  canonically conjugate to  $X$  we find

$$\varepsilon_X(\mathcal{E}_{\text{MZ}}) = \frac{1}{2}(1 - \cos\theta) \quad \text{and} \quad (14)$$

$$\nu_Z(\mathcal{E}_{\text{MZ}}) = \eta_Z(\mathcal{E}) = \widehat{\eta}_Z(\mathcal{E}) = \frac{1}{2}(1 - \sin\theta). \quad (15)$$

In all of the figures of merit, the optimal recovery map  $\mathcal{R}$  is to do nothing, while in  $\widehat{\eta}_Z$  the optimal channel  $\mathcal{C}$  outputs the average of the two outputs of  $\mathcal{P}_Z\mathcal{E}$ .

## 4 Uncertainty relations in finite dimensions

### 4.1 Complementarity measures

Before turning to the uncertainty relations, we first present several measures of complementarity that will appear therein. Indeed, we can use the above notions of disturbance to define several measures of complementarity that will later appear in our uncertainty relations. For instance, we can measure the complementarity of two observables just by using the measurement disturbance  $\nu$ . Specifically, treating  $\mathcal{Q}_X$  as the actual measurement and  $\mathcal{Q}_Z$  as the ideal measurement, we define  $c_M(X, Z) := \nu_Z(\mathcal{Q}_X)$ . This quantity is equivalent to  $\varepsilon_Z(\mathcal{Q}_X)$  since any recovery map  $\mathcal{R}_{X \rightarrow Z}$  in  $\varepsilon_Z$  can be used to define  $\mathcal{R}'_{X \rightarrow A}$  in  $\nu_Z$  by  $\mathcal{R}' = \mathcal{R}\mathcal{P}_Z$ . Similarly, we could treat one observable as defining the ideal state preparation device and the other as the measurement apparatus, which leads to  $c_P(X, Z) := \eta_Z(\mathcal{Q}_X)$ . Here we could also use the “figure of demerit” and define  $\widehat{c}_P(X, Z) := \widehat{\eta}_Z(\mathcal{Q}_X)$ .

Though the three complementarity measures are conceptually straightforward, it is also desirable to have closed-form expressions, particularly for the bounds in the uncertainty relations. To this end, we derive lower bounds as follows. First, consider  $c_M$  and choose as inputs  $Z$  basis states. This gives, for random choice of input,

$$c_M(X, Z) \geq \inf_{\mathcal{R}} \delta(\mathcal{P}_Z\mathcal{Q}_Z, \mathcal{P}_Z\mathcal{Q}_X\mathcal{R}) \quad (16a)$$

$$\geq 1 - \max_R \frac{1}{d} \sum_{xz} |\langle\varphi_x|\theta_z\rangle|^2 R_{zx} \quad (16b)$$

$$\geq 1 - \max_R \frac{1}{d} \sum_x \max_z |\langle\varphi_x|\theta_z\rangle|^2 \sum_{z'} R_{z'x} \quad (16c)$$

$$= 1 - \frac{1}{d} \sum_x \max_z |\langle\varphi_x|\theta_z\rangle|^2, \quad (16d)$$

where the maximization is over stochastic matrices  $R$ , and we use the fact that  $\sum_z R_{zx} = 1$  for all  $x$ . For  $c_P$  we can proceed similarly. Again replacing the recovery map  $\mathcal{R}_{X \rightarrow A}$  followed by  $\mathcal{Q}_Z$  with a classical postprocessing

map  $\mathcal{R}_{X \rightarrow Z}$ , we have

$$c_p(X, Z) \geq \inf_{\mathcal{R}_{X \rightarrow A}} \delta(\mathcal{P}_Z \mathcal{Q}_Z, \mathcal{P}_Z \mathcal{Q}_X \mathcal{R} \mathcal{Q}_Z) \quad (17a)$$

$$= \inf_{\mathcal{R}_{X \rightarrow Z}} \delta(\mathcal{P}_Z \mathcal{Q}_Z, \mathcal{P}_Z \mathcal{Q}_X \mathcal{R}) \quad (17b)$$

$$\geq 1 - \frac{1}{d} \sum_x \max_z |\langle \varphi_x | \theta_z \rangle|^2. \quad (17c)$$

For  $\hat{c}_p(X, Z)$  we have

$$\hat{c}_p(X, Z) = \frac{d-1}{d} - \inf_{\mathcal{C}: \text{const.}} \delta(\mathcal{C}, \mathcal{P}_Z \mathcal{Q}_X) \quad (18a)$$

$$= \frac{d-1}{d} - \min_P \max_z \delta(P, \mathcal{Q}_X^*(\theta_z)) \quad (18b)$$

$$\geq \frac{d-1}{d} - \max_z \frac{1}{2} \sum_x \left| \frac{1}{d} - |\langle \varphi_x | \theta_z \rangle|^2 \right|, \quad (18c)$$

where the bound comes from choosing  $P$  to be the uniform distribution. We could also choose  $P(x) = |\langle \varphi_x | \theta_{z'} \rangle|^2$  for some  $z'$  to obtain the bound  $\hat{c}_p(X, Z) \geq \frac{d-1}{d} - \min_{z'} \max_z \frac{1}{2} \sum_x |\text{Tr}[\varphi_x(\theta_z - \theta_{z'})]|$ . However, from numerical investigation of random bases, it appears that this bound is rarely better than the previous one.

Let us comment on the properties of the complementarity measures and their bounds in (16d), (17c), and (18c). Both expressions in the bounds are, properly, functions only of the two orthonormal bases involved, depending only on the set of overlaps. In particular, both are invariant under relabelling the bases. Uncertainty relations formulated in terms of conditional entropy typically only involve the largest overlap or largest two overlaps [7, 57], but the bounds derived here are yet more sensitive to the structure of the overlaps. Interestingly, the quantity in (16d) appears in the information exclusion relation of [57], where the sum of mutual informations different systems can have about the observables  $X$  and  $Z$  is bounded by  $\log_2 d \sum_x \max_z |\langle \varphi_x | \theta_z \rangle|^2$ .

The complementarity measures themselves all take the same value in two extreme cases: zero in the trivial case of identical bases,  $(d-1)/d$  in the case that the two bases are conjugate, meaning  $|\langle \varphi_x | \theta_z \rangle|^2 = 1/d$  for all  $x, z$ . In between, however, the separation between the two can be quite large. Consider two observables that share two eigenvectors while the remainder are conjugate. The bounds (16d) and (17c) imply that  $c_M$  and  $c_p$  are both greater than  $(d-3)/d$ . The bound on  $\hat{c}_p$  from (18c) is zero, though a better choice of constant channel can easily be found in this case. In dimensions  $d = 3k + 2$ , fix the constant channel to output the distribution  $P$  with probability  $1/3$  of being either of the last two outputs,  $1/3k$  for any  $k$  of the remainder, and zero otherwise. Then we have  $\hat{c}_p \geq \frac{d-1}{d} - \max_z \delta(P, \mathcal{Q}_X^* \mathcal{P}_Z^*(z))$ . It is easy to show the optimal value is  $2/3$  so that  $\hat{c}_p \geq (d-3)/3d$ . Hence, in the limit of large  $d$ , the gap between the two measures can be at least  $2/3$ . This example also shows that the gap between the complementary measures and the bounds can be large, though we will not investigate this further here.

## 4.2 Results

We finally have all the pieces necessary to formally state our uncertainty relations. The first relates measurement error and measurement disturbance, where we have

**Theorem 1.** *For any two observables  $X$  and  $Z$  and any quantum instrument  $\mathcal{E}$ ,*

$$\sqrt{2\varepsilon_X(\mathcal{E})} + \nu_Z(\mathcal{E}) \geq c_M(X, Z) \quad \text{and} \quad (19)$$

$$\varepsilon_X(\mathcal{E}) + \sqrt{2\nu_Z(\mathcal{E})} \geq c_M(Z, X). \quad (20)$$

Due to Lemma 1, any joint measurement of two observables can be decomposed into a sequential measurement, which implies that these bounds hold for joint measurement devices as well. Indeed, we will make use of that lemma to derive (20) from (19) in the proof below. Of course we can replace the  $c_M$  quantities with closed-form expressions using the bound in (16d). Figure 6 shows the bound for the case of conjugate observables of a qubit, for which  $c_M(X, Z) = c_M(Z, X) = \frac{1}{2}$ . It also shows the particular relation between error and measurement disturbance achieved by the apparatus  $\mathcal{E}_{MZ}$  mentioned at the end of §3, from which we can conclude that that bound is tight in the region of vanishing error or vanishing disturbance.

For measurement error and preparation disturbance we find the following relations

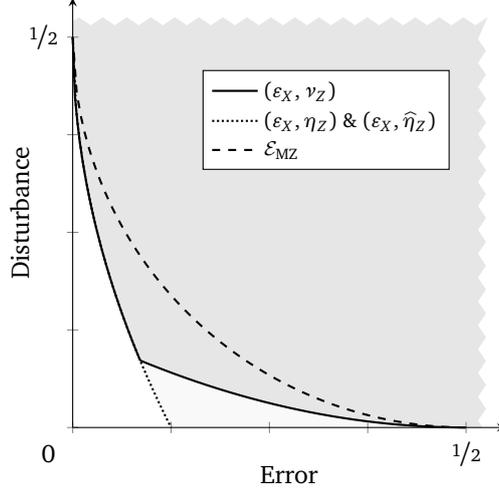


Figure 6: Error versus disturbance bounds for conjugate qubit observables. Theorem 1 restricts the possible combinations of measurement error  $\varepsilon_X$  and measurement disturbance  $\nu_Z$  to the dark gray region bounded by the solid line. Theorem 2 additionally includes the light gray region. Also shown are the error and disturbance values achieved by  $\mathcal{E}_{MZ}$  from §3.

**Theorem 2.** For any two observables  $X$  and  $Z$  and any quantum instrument  $\mathcal{E}$ ,

$$\sqrt{2\varepsilon_X(\mathcal{E})} + \eta_Z(\mathcal{E}) \geq c_P(X, Z) \quad \text{and} \quad (21)$$

$$\sqrt{2\varepsilon_X(\mathcal{E})} + \hat{\eta}_Z(\mathcal{E}) \geq \hat{c}_P(X, Z). \quad (22)$$

Returning to Figure 6 but replacing the vertical axis with  $\eta_Z$  or  $\hat{\eta}_Z$ , we now have only the upper branch of the bound, which continues to the horizontal axis as the dotted line. Here we can only conclude that the bounds are tight in the region of vanishing error.

### 4.3 Proofs

The proofs of all three uncertainty relations are just judicious applications of the triangle inequality, and the particular bound comes from the setting in which  $\mathcal{P}_Z$  meets  $\mathcal{Q}_X$ . We shall make use of the fact that an instrument which has a small error in measuring  $\mathcal{Q}_X$  is close to one which actually employs the instrument associated with  $\mathcal{Q}_X$ . This is encapsulated in the following

**Lemma 2.** For any apparatus  $\mathcal{E}_{A \rightarrow YB}$  there exists a channel  $\mathcal{F}_{XA \rightarrow YB}$  such that  $\delta(\mathcal{E}, \mathcal{Q}'_X \mathcal{F}) \leq \sqrt{2\varepsilon_X(\mathcal{E})}$ , where  $\mathcal{Q}'_X$  is a quantum instrument associated with the measurement  $\mathcal{Q}_X$ . Furthermore, if  $\mathcal{Q}_X$  is a projective measurement, then there exists a state preparation  $\mathcal{P}_{X \rightarrow YB}$  such that  $\delta(\mathcal{E}, \mathcal{Q}_X \mathcal{P}) \leq \sqrt{2\varepsilon_X(\mathcal{E})}$ .

*Proof.* Let  $V : \mathcal{H}_A \rightarrow \mathcal{H}_B \otimes \mathcal{H}_E \otimes L^2(X)$  and  $W_X : \mathcal{H}_A \rightarrow L^2(X) \otimes \mathcal{H}_A$  be respective dilations of  $\mathcal{E}$  and  $\mathcal{Q}_X$ . Using the dilation  $W_X$  we can define the instrument  $\mathcal{Q}'_X$  as

$$\begin{aligned} \mathcal{Q}'_X : L^\infty(X) \otimes \mathcal{B}(\mathcal{H}_B) &\rightarrow \mathcal{B}(\mathcal{H}_A) \\ g \otimes A &\mapsto W_X^*(\pi(g) \otimes A)W_X. \end{aligned} \quad (23)$$

Suppose  $\mathcal{R}_{Y \rightarrow X}$  is the optimal map in the definition of  $\varepsilon_X(\mathcal{E})$ , and let  $\mathcal{R}'_{Y \rightarrow XY}$  be the extension of  $\mathcal{R}$  which keeps the input  $Y$ ; it has a dilation  $V' : L^2(Y) \rightarrow L^2(Y) \otimes L^2(X)$ . By Stinespring continuity, in finite dimensions there exists a conditional isometry  $U_X : L^2(X) \otimes \mathcal{H}_A \rightarrow L^2(X) \otimes L^2(Y) \otimes \mathcal{H}_B \otimes \mathcal{H}_E$  such that

$$\|V'V - U_X W_X\|_\infty \leq \sqrt{2\varepsilon_X(\mathcal{E})}. \quad (24)$$

Now consider the map

$$\begin{aligned} \mathcal{E}' : L^\infty(Y) \otimes \mathcal{B}(\mathcal{H}_B) &\rightarrow \mathcal{B}(\mathcal{H}_A) \\ f \otimes A &\mapsto W_X^* U_X^* (\mathbb{1}_X \otimes \pi(f) \otimes A \otimes \mathbb{1}_E) U_X W_X. \end{aligned} \quad (25)$$

By the other bound in Stinespring continuity we thus have  $\delta(\mathcal{E}, \mathcal{E}') \leq \sqrt{2\varepsilon_X(\mathcal{E})}$ . Furthermore, as described in §2.4,  $U_X$  is a conditional isometry, i.e. a collection of isometries  $U_x : \mathcal{H}_A \rightarrow L^2(Y) \otimes \mathcal{H}_B \otimes \mathcal{H}_E$  for each measurement outcome  $x$ . Note that we may regard elements of  $L^\infty(X) \otimes \mathcal{B}(\mathcal{H})$  as sequences  $(A_x)_{x \in X}$  with  $A_x \in \mathcal{B}(\mathcal{H})$  for all  $x \in X$  such that  $\text{ess sup}_x \|A_x\|_\infty < \infty$ . Therefore we may define

$$\begin{aligned} \mathcal{F} : L^\infty(Y) \otimes \mathcal{B}(\mathcal{H}_B) &\rightarrow L^\infty(X) \otimes \mathcal{B}(\mathcal{H}_A) \\ f \otimes A &\mapsto (U_x^*(\pi(f) \otimes A \otimes \mathbb{1}_E)U_x)_{x \in X}, \end{aligned} \quad (26)$$

so that  $\mathcal{E}' = Q'_X \mathcal{F}$ . This completes the proof of the first statement.

If  $Q_X$  is a projective measurement, then the output  $B$  of  $Q'_X$  can just as well be prepared from the  $X$  output. Describing this with the map  $\mathcal{P}'_{X \rightarrow XA}$  which prepares states in  $A$  given the value of  $X$  and retains  $X$  at the output, we have  $Q'_X = Q_X \mathcal{P}'$ . Setting  $\mathcal{P} = \mathcal{P}' \mathcal{F}$  completes the proof of the second statement.  $\square$

Now, to prove (19), start with the triangle inequality and monotonicity. Suppose  $\mathcal{P}_{X \rightarrow YB}$  is the state preparation map from Lemma 2. Then, for any  $\mathcal{R}_{YB \rightarrow A}$ ,

$$\delta(Q_Z, Q_X \mathcal{P} \mathcal{R} Q_Z) \leq \delta(Q_Z, \mathcal{E} \mathcal{R} Q_Z) + \delta(\mathcal{E} \mathcal{R} Q_Z, Q_X \mathcal{P} \mathcal{R} Q_Z) \quad (27a)$$

$$\leq \delta(Q_Z, \mathcal{E} \mathcal{R} Q_Z) + \delta(\mathcal{E}, Q_X \mathcal{P}) \quad (27b)$$

$$= \delta(Q_Z, \mathcal{E} \mathcal{R} Q_Z) + \sqrt{2\varepsilon_X(\mathcal{E})}. \quad (27c)$$

Observe that  $\mathcal{P} \mathcal{R} Q_Z$  is just a map  $\mathcal{R}'_{X \rightarrow Z}$ . Taking the infimum over  $\mathcal{R}$  we then have

$$\sqrt{2\varepsilon_X(\mathcal{E})} + \nu_Z(\mathcal{E}) \geq \inf_{\mathcal{R}} \delta(Q_Z, Q_X \mathcal{P} \mathcal{R} Q_Z) \quad (28a)$$

$$\geq \inf_{\mathcal{R}} \delta(Q_Z, Q_X \mathcal{R}). \quad (28b)$$

To show (20), let  $\mathcal{R}_{YB \rightarrow A}$  and  $\mathcal{R}'_{Y \rightarrow X}$  be the optimal maps in  $\nu_Z(\mathcal{E})$  and  $\varepsilon_X(\mathcal{E})$ , respectively. Now apply Lemma 1 to  $\mathcal{M} = \mathcal{E} \mathcal{R}' \mathcal{R} Q_Z$  and suppose that  $\mathcal{E}'_{A \rightarrow ZB}$  is the resulting instrument and  $\mathcal{M}_{ZB \rightarrow X}$  is the conditional measurement. By the above argument,  $\sqrt{2\varepsilon_Z(\mathcal{E}')} + \nu_X(\mathcal{E}') \geq \inf_{\mathcal{R}} \delta(Q_X, Q_Z \mathcal{R})$ . But  $\varepsilon_Z(\mathcal{E}') \leq \delta(Q_Z, \mathcal{E}' \mathcal{T}_B) = \nu_Z(\mathcal{E})$  and  $\nu_X(\mathcal{E}') \leq \delta(Q_X, \mathcal{E}' \mathcal{M}) = \varepsilon_X(\mathcal{E})$ , where in the latter we use the fact that we could always reprepare an  $X$  eigenstate and then let  $Q_X$  measure it. Therefore the desired bound holds.

To establish (21), we proceed just as above to obtain

$$\delta(\mathcal{P}_Z, \mathcal{P}_Z Q_X \mathcal{P} \mathcal{R}) \leq \delta(\mathcal{P}_Z, \mathcal{P}_Z \mathcal{E} \mathcal{R}) + \sqrt{2\varepsilon_X(\mathcal{E})}. \quad (29)$$

Now  $\mathcal{P}_{X \rightarrow YB} \mathcal{R}_{YB \rightarrow A}$  is a preparation map  $\mathcal{P}_{X \rightarrow A}$ , and taking the infimum over  $\mathcal{R}$  gives

$$\sqrt{2\varepsilon_X(\mathcal{E})} + \eta_Z(\mathcal{E}) \geq \inf_{\mathcal{R}} \delta(\mathcal{P}_Z, \mathcal{P}_Z Q_X \mathcal{P} \mathcal{R}) \quad (30a)$$

$$\geq \inf_{\mathcal{P}} \delta(\mathcal{P}_Z, \mathcal{P}_Z Q_X \mathcal{P}). \quad (30b)$$

Finally, (22). Since the  $\hat{\eta}_Z$  disturbance measure is defined “backwards”, we start the triangle inequality with the distinguishability quantity related to disturbance, rather than the eventual constant of the bound. For any channel  $\mathcal{C}_{Z \rightarrow X}$  and  $\mathcal{P}_{X \rightarrow YB}$  from Lemma 2, just as before we have

$$\delta(\mathcal{C} \mathcal{P}, \mathcal{P}_Z \mathcal{E}) \leq \delta(\mathcal{C} \mathcal{P}, \mathcal{P}_Z Q_X \mathcal{P}) + \delta(\mathcal{P}_Z Q_X \mathcal{P}, \mathcal{P}_Z \mathcal{E}) \quad (31a)$$

$$\leq \delta(\mathcal{C}, \mathcal{P}_Z Q_X) + \sqrt{2\varepsilon_X(\mathcal{E})}. \quad (31b)$$

Now we take the infimum over constant channels  $\mathcal{C}_{Z \rightarrow X}$ . Note that

$$\inf_{\mathcal{C}_{Z \rightarrow YB}} \delta(\mathcal{C}, \mathcal{P}_Z \mathcal{E}) \leq \inf_{\mathcal{C}_{Z \rightarrow X}} \delta(\mathcal{C} \mathcal{P}, \mathcal{P}_Z \mathcal{E}). \quad (32)$$

Therefore, we have

$$\sqrt{2\varepsilon_X(\mathcal{E})} + \hat{\eta}_Z(\mathcal{E}) \geq \frac{d-1}{d} - \inf_{\mathcal{C}} \delta(\mathcal{C}, \mathcal{P}_Z Q_X). \quad (33)$$

This last proof also applies to a more general definition of disturbance which does not use  $\mathcal{P}_Z$  at the input, but rather diagonalizes or “pinches” any input quantum system in the  $Z$  basis. Such a transformation can

be thought of as the result of performing an ideal  $Z$  measurement, but forgetting the result. More formally, letting  $\mathcal{Q}_Z^{\natural} = \mathcal{W}_Z \mathcal{T}_Z$  with  $\mathcal{W}_Z : a \rightarrow W_Z^* a W_Z$ , we can define

$$\tilde{\eta}_Z(\mathcal{E}) = \frac{d-1}{d} - \inf_C \delta(\mathcal{C}, \mathcal{Q}_Z^{\natural} \mathcal{E}). \quad (34)$$

Though perhaps less conceptually appealing, this is a more general notion of disturbance, since now we can potentially use entanglement at the input to increase distinguishability of  $\mathcal{Q}_Z^{\natural} \mathcal{E}$  from any constant channel. However, due to the form of  $\mathcal{Q}_Z^{\natural}$ , entanglement will not help. Applied to any bipartite state, the map  $\mathcal{Q}_Z^{\natural}$  produces a state of the form  $\sum_z p_z |\theta_z\rangle \langle \theta_z| \otimes \sigma_z$  for some probability distribution  $p_z$  and set of normalized states  $\sigma_z$ , and therefore the input to  $\mathcal{E}$  itself is again an output of  $\mathcal{P}_Z$ . Since classical correlation with ancillary systems is already covered in  $\hat{\eta}_Z(\mathcal{E})$ , it follows that  $\tilde{\eta}_Z(\mathcal{E}) = \hat{\eta}_Z(\mathcal{E})$ .

## 5 Position & momentum

### 5.1 Gaussian precision-limited measurement and preparation

Now we turn to the infinite-dimensional case of position and momentum measurements. Let us focus on Gaussian limits on precision, where the convolution function  $\alpha$  described in §2.2 is the square root of a normalized Gaussian of width  $\sigma$ , and for convenience define

$$g_{\sigma}(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}}. \quad (35)$$

One advantage of the Gaussian choice is that the Stinespring dilation of the ideal  $\sigma$ -limited measurement device is just a canonical transformation. Thus, measurement of position  $Q$  just amounts to adding this value to an ancillary system which is prepared in a zero-mean Gaussian state with position standard deviation  $\sigma_Q$ , and similarly for momentum. The same interpretation is available for precision-limited state preparation. To prepare a momentum state of width  $\sigma_p$ , we begin with a system in a zero-mean Gaussian state with momentum standard deviation  $\sigma_p$  and simply shift the momentum by the desired amount.

Given the ideal devices, the definitions of error and disturbance are those of §3, as in the finite-dimensional case, with the slight change that the first term of  $\hat{\eta}$  is now 1. To reduce clutter, we do not indicate  $\sigma_Q$  and  $\sigma_p$  specifically in the error and disturbance functions themselves.

Since our error and disturbance measures are based on possible state preparations and measurements in order to best distinguish the two devices, in principle one ought to consider precision limits in the distinguishability quantity  $\delta$ . However, we will not follow this approach here, and instead we allow test of arbitrary precision in order to preserve the link between distinguishability and the cb norm. This leads to bounds that are perhaps overly pessimistic, but nevertheless limit the possible performance of any device.

### 5.2 Results

As discussed previously, the disturbance measure of demerit  $\hat{\eta}$  cannot be expected to lead to uncertainty relations for position and momentum observables, as any non-constant channel can be perfectly differentiated from a constant one by inputting states of arbitrarily high momentum. We thus focus on the disturbance measures of merit.

**Theorem 3.** *Set  $c = 2\sigma_Q\sigma_p$  for any precision values  $\sigma_Q, \sigma_p > 0$ . Then for any quantum instrument  $\mathcal{E}$ ,*

$$\left. \begin{aligned} \sqrt{2\varepsilon_Q(\mathcal{E})} + \nu_p(\mathcal{E}) \\ \varepsilon_Q(\mathcal{E}) + \sqrt{2\nu_Q(\mathcal{E})} \end{aligned} \right\} \geq \frac{1-c^2}{(1+c^{2/3}+c^{4/3})^{3/2}} \quad \text{and} \quad (36)$$

$$\sqrt{2\varepsilon_Q(\mathcal{E})} + \eta_p(\mathcal{E}) \geq \frac{(1+c^2)^{1/2}}{((1+c^2)+c^{2/3}(1+c^2)^{2/3}+c^{4/3}(1+c^2)^{1/3})^{3/2}}. \quad (37)$$

Before proceeding to the proofs, let us comment on the properties of the two bounds. As can be seen in Figure 7, the bounds take essentially the same values for  $\sigma_Q\sigma_p \ll \frac{1}{2}$ , and indeed both evaluate to unity at  $\sigma_Q\sigma_p = 0$ . This is the region of combined position and momentum precision far smaller than the natural scale set by  $\hbar$ , and the limit of infinite precision accords with the finite-dimensional bounds for conjugate observables in the limit  $d \rightarrow \infty$ . Otherwise, though, the bounds differ remarkably. The measurement disturbance bound in (36) is positive only when  $\sigma_Q\sigma_p \leq \frac{1}{2}$ , which is the Heisenberg precision limit. In contrast, the preparation disturbance bound in (37) is always positive, though it decays roughly as  $(\sigma_Q\sigma_p)^2$ .

The distinction between these two cases is a result of allowing arbitrarily precise measurements in the distinguishability measure. It can be understood by the following heuristic argument. Consider an experiment in which a momentum state of width  $\sigma_p^{\text{in}}$  is subjected to a position measurement of resolution  $\sigma_Q$  and then a momentum measurement of resolution  $\sigma_p^{\text{out}}$ . From the uncertainty principle, we expect the position measurement to change the momentum by an amount  $\sim 1/\sigma_Q$ . Thus, to reliably detect the change in momentum,  $\sigma_p^{\text{out}}$  must fulfill the condition  $\sigma_p^{\text{out}} \ll \sigma_p^{\text{in}} + 1/\sigma_Q$ . The Heisenberg limit in the measurement disturbance scenario is  $\sigma_p^{\text{out}} = 2/\sigma_Q$ , meaning this condition cannot be met no matter how small we choose  $\sigma_p^{\text{in}}$ . This is consistent with no nontrivial bound in (36) in this region. On the other hand, for preparation disturbance the Heisenberg limit is  $\sigma_p^{\text{in}} = 2/\sigma_Q$ , so detecting the change in momentum simply requires  $\sigma_p^{\text{out}} \ll 1/\sigma_Q$ . A more satisfying approach would be to include the precision limitation in the distinguishability measure to restore the symmetry of the two scenarios, but this requires significant changes to the proof and is left for future work.

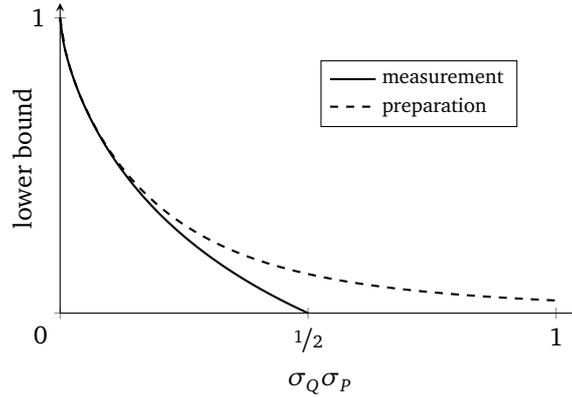


Figure 7: Uncertainty bounds appearing in Theorem 3 in terms of the combined precision  $\sigma_Q \sigma_p$ . The solid line corresponds to the bound involving measurement disturbance, (36), the dashed line to the bound involving preparation disturbance, (37).

### 5.3 Proofs

The proof of Theorem 3 is broadly similar to the finite-dimensional case. We would again like to begin with  $\mathcal{F}_{Q_A \rightarrow Y_B}$  from Lemma 2 such that  $\delta(\mathcal{E}, \mathcal{Q}'_Q \mathcal{F}) \leq \sqrt{2\varepsilon_Q(\mathcal{E})}$ . However, the argument does not quite go through, as in infinite dimensions we cannot immediately ensure that the infimum in Stinespring continuity is attained. Nonetheless, we can consider a sequence of maps  $(\mathcal{F}_n)_{n \in \mathbb{N}}$  such that the desired distinguishability bound holds in the limit  $n \rightarrow \infty$ .

To show (36), we follow the steps in (27). Now, though, consider the map  $\mathcal{F}'_n$  which just appends  $Q$  to the output of  $\mathcal{F}_n$ , and define  $\mathcal{N} = \mathcal{Q}'_Q \mathcal{F}_n \mathcal{R} \mathcal{Q}_p$ , where  $\mathcal{Q}'_Q$  is the instrument associated with position measurement  $\mathcal{Q}_Q$ . Then we have

$$\delta(\mathcal{Q}_p, \mathcal{N} \mathcal{T}_Q) \leq \delta(\mathcal{Q}_p, \mathcal{E} \mathcal{R} \mathcal{Q}_p) + \delta(\mathcal{E} \mathcal{R} \mathcal{Q}_p, \mathcal{N} \mathcal{T}_Q) \quad (38a)$$

$$\leq \delta(\mathcal{Q}_p, \mathcal{E} \mathcal{R} \mathcal{Q}_p) + \delta(\mathcal{E}, \mathcal{Q}'_Q \mathcal{F}_n). \quad (38b)$$

Taking the limit  $n \rightarrow \infty$  and the infimum over recovery maps  $\mathcal{R}$  produces  $\sqrt{2\varepsilon_Q(\mathcal{E})} + \nu_p(\mathcal{E})$  on the righthand side. We can bound the lefthand side by testing with pure unentangled inputs:

$$\delta(\mathcal{Q}_p, \mathcal{N} \mathcal{T}_Q) \geq \sup_{\psi, f} \langle \psi, (\mathcal{Q}_p(f) - [\mathcal{N} \mathcal{T}_Q](f)) \psi \rangle. \quad (39)$$

Now we want to show that, since  $\mathcal{Q}_p$  is covariant with respect to phase space translations, without loss of generality we can take  $\mathcal{N}$  to be covariant as well. Consider the translated version of both  $\mathcal{Q}_p$  and  $\mathcal{N} \mathcal{T}_Q$ , obtained by shifting their inputs and outputs correspondingly by some amount  $z = (q, p)$ . For the states  $\psi$  this shift is implemented by the Weyl-Heisenberg operators  $V_z$ , while for tests  $f$  only the value of  $p$  is relevant. Any such shift does not change the distinguishability, because we can always shift  $\psi$  and  $f$  as well to recover the original quantity. Averaging over the translated versions therefore also leads to the same distinguishability, and since  $\mathcal{Q}_p$  is itself covariant, the averaging results in a covariant  $\mathcal{N} \mathcal{T}_Q$ . The details of the averaging require some care in this noncompact setting, but are standard by now, and we refer the reader

to the work of Werner [22] for further details. Since  $\mathcal{T}_Q$  just ignores the Q output of the measurement  $\mathcal{N}$ , we may thus proceed by assuming that  $\mathcal{N}$  is a covariant measurement.

Any covariant  $\mathcal{N}$  has the form

$$\mathcal{N}(f) = \int_{\mathbb{R}^2} \frac{dz}{2\pi} f(z) V_z m V_z^*, \quad (40)$$

for some positive operator  $m$  such that  $\text{Tr}[m] = 1$ . Due to the definition of  $\mathcal{N}$ , the position measurement result is precisely that obtained from  $\mathcal{Q}_Q$ . By the covariant form of  $\mathcal{N}$ , this implies that the position width of  $m$  is just  $\sigma_Q$  (or rather that of the parity version of  $m$ , see [22]). Suppose the momentum distribution has standard deviation  $\hat{\sigma}_p$ ; then  $\sigma_Q \hat{\sigma}_p \geq 1/2$  follows from the Kennard uncertainty relation [3].

Now we can evaluate the lower bound term by term. Let us choose a Gaussian state in the momentum representation and test function:  $\psi = g_{\sigma_\psi}^{\frac{1}{2}}$  and  $f = \sqrt{2\pi} \sigma_f g_{\sigma_f}$ . Then the first term is a straightforward Gaussian integral, since the precision-limited measurement just amounts to the ideal measurement convolved with  $g_{\sigma_p}$ :

$$\langle \psi, \mathcal{Q}_p(f) \psi \rangle = \int_{\mathbb{R}^2} dp' dp g_{\sigma_\psi}(p') g_{\sigma_p}(p' - p) f(p) \quad (41a)$$

$$= \frac{\sigma_f}{\sqrt{\sigma_f^2 + \sigma_p^2 + \sigma_\psi^2}}. \quad (41b)$$

The second term is the same, just with  $\hat{\sigma}_p$  instead of  $\sigma_p$ , so we have

$$\delta(\mathcal{Q}_p, \mathcal{N} \mathcal{T}_Q) \geq \frac{\sigma_f}{\sqrt{\sigma_f^2 + \sigma_p^2 + \sigma_\psi^2}} - \frac{\sigma_f}{\sqrt{\sigma_f^2 + \hat{\sigma}_p^2 + \sigma_\psi^2}}. \quad (42)$$

The tightest possible bound comes from the smallest  $\hat{\sigma}_p$ , which is  $1/2\sigma_Q$ , and the bound is clearly trivial if  $\sigma_Q \sigma_p \geq 1/2$ . If this is not the case, we can optimize our choice of  $\sigma_f$ . To simplify the calculation, assume that  $\sigma_\psi$  is small compared to  $\sigma_f$  (so that we are testing with a very narrow momentum state). Then, with  $c = 2\sigma_Q \sigma_p$ , the optimal  $\sigma_f$  is given by

$$\sigma_f^2 = \frac{\sigma_p^2}{c^{2/3}(1 + c^{2/3})}. \quad (43)$$

Using this in (42) gives (36).

For preparation disturbance, proceed as before to obtain

$$\delta(\mathcal{P}_p, \mathcal{P}_p \mathcal{Q}'_Q \mathcal{F}'_n \mathcal{R} \mathcal{T}_Q) \leq \delta(\mathcal{P}_p, \mathcal{P}_p \mathcal{E} \mathcal{R}) + \delta(\mathcal{P}_p \mathcal{E} \mathcal{R}, \mathcal{P}_p \mathcal{Q}'_Q \mathcal{F}'_n \mathcal{R} \mathcal{T}_Q) \quad (44a)$$

$$\leq \delta(\mathcal{P}_p, \mathcal{P}_p \mathcal{E} \mathcal{R}) + \delta(\mathcal{E}, \mathcal{Q}'_Q \mathcal{F}_n) \quad (44b)$$

Now the limit  $n \rightarrow \infty$  and the infimum over recovery maps  $\mathcal{R}$  produces  $\sqrt{2\varepsilon_Q(\mathcal{E})} + \eta_p(\mathcal{E})$  on the righthand side. A lower bound on the quantity on the lefthand side can be obtained by using  $\mathcal{P}_p$  to prepare a  $\sigma_p$ -limited input state and making a  $\sigma_m$ -limited momentum measurement  $\mathcal{Q}_p$  measurement on the output, so that, for  $\mathcal{N}$  as before,

$$\delta(\mathcal{P}_p, \mathcal{P}_p \mathcal{Q}'_Q \mathcal{F}'_n \mathcal{R} \mathcal{T}_Q) \geq \sup_{\psi: \text{Gaussian}; f} \langle \psi, (\mathcal{Q}_p(f) - [\mathcal{N} \mathcal{T}_Q](f)) \psi \rangle. \quad (45)$$

The only difference to (39) is that the supremum is restricted to Gaussian states of width  $\sigma_p$ . The covariance argument nonetheless goes through as before, and we can proceed to evaluate the lower bound as above. This yields

$$\delta(\mathcal{P}_p, \mathcal{P}_p \mathcal{Q}'_Q \mathcal{F}'_n \mathcal{R} \mathcal{T}_Q) \geq \frac{\sigma_f}{\sqrt{\sigma_f^2 + \sigma_m^2 + \sigma_p^2}} - \frac{\sigma_f}{\sqrt{\sigma_f^2 + \frac{1}{4\sigma_Q^2} + \sigma_p^2}}. \quad (46)$$

We may as well consider  $\sigma_m \rightarrow 0$  so as to increase the first term. The optimal  $\sigma_f$  is then given by the optimizer above, replacing  $c$  with  $c/\sqrt{1 + c^2}$ . Making the same replacement in (36) yields (37).

## 6 Applications

### 6.1 No information about $Z$ without disturbance to $X$

A useful tool in the construction of quantum information processing protocols is the link between reliable transmission of  $X$  eigenstates through a channel  $\mathcal{N}$  and  $Z$  eigenstates through its complement  $\mathcal{N}^\sharp$ , particularly when the observables  $X$  and  $Z$  are maximally complementary, i.e.  $|\langle \varphi_x | \vartheta_z \rangle|^2 = \frac{1}{d}$  for all  $x, z$ . Due to the uncertainty principle, we expect that a channel cannot reliably transmit the bases to different outputs, since this would provide a means to simultaneously measure  $X$  and  $Z$ . This link has been used by Shor and Preskill to prove the security of quantum key distribution [58] and by Devetak to determine the quantum channel capacity [59]. Entropic state-preparation uncertainty relations from [6, 44] can be used to understand both results, as shown in [60, 61].

However, the above approach has the serious drawback that it can only be used in cases where the specific  $X$ -basis transmission over  $\mathcal{N}$  and  $Z$ -basis transmission over  $\mathcal{N}^\sharp$  are in some sense compatible and not *counterfactual*; because the argument relies on a state-dependent uncertainty principle, both scenarios must be compatible with the same quantum state. Fortunately, this can be done for both QKD security and quantum capacity, because at issue is whether  $X$ -basis ( $Z$ -basis) transmission is reliable (unreliable) on average *when the states are selected uniformly at random*. Choosing among either basis states at random is compatible with a random measurement in either basis of half of a maximally entangled state, and so both  $X$  and  $Z$  basis scenarios are indeed compatible. The same restriction to choosing input states uniformly appears in the recent result of [33], as it also ultimately relies on a state-preparation uncertainty relation.

Using Theorem 2 we can extend the method above to counterfactual uses of arbitrary channels  $\mathcal{N}$ , in the following sense: If acting with the channel  $\mathcal{N}$  does not substantially affect the possibility of performing an  $X$  measurement, then  $Z$ -basis inputs to  $\mathcal{N}^\sharp$  result in an essentially constant output. More concretely, we have

**Corollary 1.** *Given a channel  $\mathcal{N}$  and complementary channel  $\mathcal{N}^\sharp$ , suppose that there exists a measurement  $\Lambda_X$  such that  $\delta(\mathcal{Q}_X, \mathcal{N}\Lambda_X) \leq \varepsilon$ . Then there exists a constant channel  $\mathcal{C}$  such that*

$$\delta(\mathcal{Q}_Z^\sharp \mathcal{N}^\sharp, \mathcal{C}) \leq \sqrt{2\varepsilon} + \frac{d-1}{d} - \widehat{c}_p(X, Z). \quad (47)$$

For maximally complementary  $X$  and  $Z$ ,  $\delta(\mathcal{Q}_Z^\sharp \mathcal{N}^\sharp, \mathcal{C}) \leq \sqrt{2\varepsilon}$ .

*Proof.* Let  $V$  be the Stinespring dilation of  $\mathcal{N}$  such that  $\mathcal{N}^\sharp$  is the complementary channel and define  $\mathcal{E} = \mathcal{V}_{\mathcal{N}}\Lambda_X$ . For  $\mathcal{C}$  the optimal choice in the definition of  $\widehat{\eta}_Z(\mathcal{E})$ , (22), (34), and  $\widehat{\eta}_Z = \widehat{\eta}_Z$  imply  $\delta(\mathcal{Q}_Z^\sharp \mathcal{E}, \mathcal{C}) \leq \sqrt{2\varepsilon} + \frac{d-1}{d} - \widehat{c}_p(X, Z)$ . Since  $\mathcal{N}^\sharp$  is obtained from  $\mathcal{E}$  by ignoring the  $\Lambda_X$  measurement result,  $\delta(\mathcal{Q}_Z^\sharp \mathcal{N}^\sharp, \mathcal{C}) \leq \delta(\mathcal{Q}_Z^\sharp \mathcal{E}, \mathcal{C})$ .  $\square$

This formulation is important because in more general cryptographic and communication scenarios we are interested in the worst-case behavior of the protocol, not the average case under some particular probability distribution. For instance, in [46] the goal is to construct a classical computer resilient to leakage of  $Z$ -basis information by establishing that reliable  $X$  basis measurement is possible despite the interference of the eavesdropper. However, such an  $X$  measurement is entirely counterfactual and cannot be reconciled with the actual  $Z$ -basis usage, as the  $Z$ -basis states will be chosen *deterministically* in the classical computer.

It is important to point out that, unfortunately, calibration testing is in general completely insufficient to establish a small value of  $\delta(\mathcal{Q}_X, \mathcal{N}\Lambda_X)$ . More specifically, the following example shows that there is no dimension-independent bound connecting  $\inf_{\Lambda_X} \delta(\mathcal{Q}_X, \mathcal{N}\Lambda_X)$  to the worst case probability of incorrectly identifying an  $X$  eigenstate input to  $\mathcal{N}$ , for arbitrary  $\mathcal{N}$ . Let the quantities  $p_{yz}$  be given by  $p_{y,0} = 2/d$  for  $y = 0, \dots, d/2 - 1$ ,  $p_{y,1} = 2/d$  for  $y = d/2, \dots, d - 1$ , and  $p_{y,z} = 1/d$  otherwise, where we assume  $d$  is even, and then define the isometry  $V : \mathcal{H}_A \rightarrow \mathcal{H}_B \otimes \mathcal{H}_C \otimes \mathcal{H}_D$  as the map taking  $|z\rangle_A$  to  $\sum_y \sqrt{p_{yz}} |y\rangle_B |z\rangle_C |y\rangle_D$ . Finally, let  $\mathcal{N} : \mathcal{B}(\mathcal{H}_B) \otimes \mathcal{B}(\mathcal{H}_C) \rightarrow \mathcal{B}(\mathcal{H}_A)$  be the channel obtained by ignoring  $D$ , i.e. in the Schrödinger picture  $\mathcal{N}^*(\varrho) = \text{Tr}_D[V\varrho V^*]$ . Now consider inputs in the  $X$  basis, with  $X$  canonically conjugate to  $Z$ . As shown in Appendix C, the probability of correctly determining any particular  $X$  input is the same for all values, and is equal to  $\frac{1}{d^2} \sum_y \left( \sum_z \sqrt{p_{y,z}} \right)^2 = (d + \sqrt{2} - 2)^2/d^2$ . The worst case  $X$  error probability therefore tends to zero like  $1/d$  as  $d \rightarrow \infty$ . On the other hand,  $Z$ -basis inputs 0 and 1 to the complementary channel  $\mathcal{E}^\sharp$  result in completely disjoint output states due to the form of  $p_{yz}$ . Thus, if we consider a test which inputs one of these randomly and checks for agreement at the output, we find  $\inf_C \delta(\mathcal{Q}_Z^\sharp \mathcal{N}^\sharp, \mathcal{C}) \geq \frac{1}{2}$ . Using the bound above, this implies  $\inf_{\Lambda_X} \delta(\mathcal{Q}_X, \mathcal{N}\Lambda_X) \geq \frac{1}{8}$ . This is not 1, but the point is it is bounded away from zero and

independent of  $d$ : There must be a factor of  $d$  when converting between the worst case error probability and the distinguishability.

We can appreciate the failure of calibration in this example from a different point of view, by appealing to the information-disturbance tradeoff of [40]. Since  $\mathcal{N}$  transmits  $Z$  eigenstates perfectly to  $BC$  and  $X$  eigenstates almost perfectly, we might be tempted to conclude that the channel is close to the identity channel. However, the information-disturbance tradeoff implies that complements of channels close to the identity are close to constant channels. Clearly this is not the case here, since  $\mathcal{N}^*(|0\rangle\langle 0|)$  is distinguishable from  $\mathcal{N}^*(|1\rangle\langle 1|)$ . This point is discussed further by one of us in [62]. The counterexample constructed above is not symmetric for  $Z$  inputs, and it is an open question if calibration is sufficient in the symmetric case. For channels that are covariant with respect to the Weyl-Heisenberg group (also known as the generalized Pauli group), it is not hard to show that calibration is in fact sufficient.

## 6.2 Connection to wave-particle duality relations

In [42] Englert presents a wave-particle complementarity relation in a Mach-Zehnder interferometer, quantifying the extent to which “the observations of an interference pattern and the acquisition of which-way information are mutually exclusive”. The particle-like “which-way” information is obtained by additional detectors in the arms of the interferometer, while fringe visibility is measured by the population difference between the two output ports of the interferometer. The detectors can be thought of as producing different states in an ancilla system, depending on the path taken by the light. Englert shows the following tradeoff between the visibility  $V$  and distinguishability  $D$  of the which-way detector states:

$$V^2 + D^2 \leq 1. \quad (48)$$

We may regard the entire interferometer plus which-way detector as an apparatus  $\mathcal{E}_{MZ}$  with quantum and classical output. It turns out that  $\mathcal{E}_{MZ}$  is precisely the nonideal qubit  $X$  measurement considered in §3 and that path distinguishability is related error of  $X$  and visibility to disturbance (all of which are equal in this case by (15)) of a conjugate observable  $Z$ . More specifically, as shown in Appendix D,

$$\varepsilon_X(\mathcal{E}_{MZ}) = \frac{1}{2}(1 - D) \quad \text{and} \quad \nu_Z(\mathcal{E}_{MZ}) = \eta_Z(\mathcal{E}_{MZ}) = \hat{\eta}_Z(\mathcal{E}_{MZ}) = \frac{1}{2}(1 - V). \quad (49)$$

Therefore, (48) is also an error-preparation disturbance relation. By the same token, the uncertainty relations in Theorems 1 and 2 imply wave-particle duality relations.

Let us comment on other connections between uncertainty and duality relations. Recently, [63] showed a relation between wave-particle duality relations and entropic uncertainty relations. As discussed above, the latter are state-dependent state-preparation relations, and so the interpretation of the wave-particle duality relation is somewhat different. Here we have shown that Englert’s relation can actually be understood as a state-independent relation.

Each of the disturbance measures are related to visibility in Englert’s setup. It is an interesting question to consider a multipath interferometer to settle the question of which disturbance measure should be associated to visibility in general. From the discussion of [64], it would appear that visibility ought to be related to measurement disturbance  $\nu_Z$ , but we leave a complete analysis to future work.

## 7 Comparison to previous work

Broadly speaking, there are two main kinds of uncertainty relations: those which are constraints on fixed experiments, including the details of the input quantum state, and those that are constraints on quantum devices themselves, independent of the particular input. All of our relations are of the latter type, in contrast to entropic relations, which are typically of the former type. At a formal level, this distinction appears in whether or not the quantities involved in the precise relation depend on the input state or not.<sup>2</sup> Each type of relation certainly has its use, though when considering error-disturbance uncertainty relations, we argued in the introduction that the conceptual underpinnings of state-dependent relations describing fixed experiments are unclear. Indeed, it is precisely because of the uncertainty principle that trouble arises in defining error and disturbance in this case. Worse still, there can be no nontrivial bound relating error and disturbance which applies universally, i.e. to all states [65].

Independent of the previous question, another major contrast between different kinds of uncertainty relations is whether they depend on the values taken by the observables, or only the configuration of their eigenstates. Again, our relations are all of the latter type, but now we share this property with entropic relations. That is not to say that the observable values are completely irrelevant in our setting, merely that they

<sup>2</sup>This is separate from the issue of whether the bound depends on the state, as for instance in the Robertson relation [4].

are not necessarily relevant. In distinguishing the outputs of an ideal position measurement of given precision from the outputs of the actual device, one may indeed make use of the difference in measurement values. But this need not be the only kind of comparison.

In the recent work of Busch, Lahti, and Werner [25], the authors used the Wasserstein metric of order two, corresponding to the mean squared error, as the underlying distance  $D(.,.)$  to measure the closeness of probability distributions. If  $\mathcal{M}^Q, \mathcal{M}^P$  are the marginals of some joint measurement of position  $Q$  and momentum  $P$ , and  $X_\varrho$  denotes the distribution coming from applying the measurement  $X$  to the state  $\varrho$ , their relation reads

$$\sup_{\varrho} D(\mathcal{M}_{\varrho}^Q, Q_{\varrho}) \cdot \sup_{\varrho} D(\mathcal{M}_{\varrho}^P, P_{\varrho}) \geq c, \quad (50)$$

for some universal constant  $c$ . In [27], the authors generalize their results to arbitrary Wasserstein metrics. As in our case, the two distinguishability quantities in (50) are separately maximized over all states, and hence the resulting expression characterizes the goodness of the approximate measurement.

One could instead ask for a “coupled optimization”, a relation of the form

$$\sup_{\varrho} [D(\mathcal{M}_{\varrho}^Q, Q_{\varrho}) D(\mathcal{M}_{\varrho}^P, P_{\varrho})] \geq c', \quad (51)$$

for some other constant  $c'$ .<sup>3</sup> This approach is taken in [66] for the question of joint measurability. While this statement certainly tells us that no device can accurately measure both position and momentum for all input states, the bound  $c'$  only holds (and can only hold) for the worst possible input state. In contrast, our bounds, as well as in (50) are state-independent in the sense that the bound holds for all states. Indeed, the two approaches are more distinct than the similarities between (50) and (51) would suggest. By optimizing over input states separately, our results and those of [22, 25, 27] are statements about the properties of measurement devices themselves, independent of any particular experimental setup. State-dependent settings capture the behavior of measurement devices in specific experimental setups and must therefore account for the details of the input state.

The same set of authors also studied the case of finite-dimensional systems, in particular qubit systems, again using the Wasserstein metric of order two [26]. Their results for this case are similar, with the product in (50) replaced by a sum. Perhaps most closely related to our results is the recent work by Ipsen [34], who uses the variational distance as the underlying distinguishability measure to derive similar additive uncertainty relations. We note, however, that both [26] and [34] only consider joint measurability and do not consider the change to the state after the approximate measurement is performed, as it is done in our error-disturbance relation. Furthermore, both base their distinguishability measures on the measurement statistics of the devices alone. But this does not necessarily tell us how distinguishable two devices ultimately are, as we could employ input states entangled with ancilla systems to test them. These two measures can be different [51], even for entanglement-breaking channels [67]. In Appendix A we give an example which shows that this is also true of quantum measurements, a specific kind of entanglement-breaking channel.

Entropic quantities are another means of comparing two probability distributions, an approach taken recently by Buscemi *et al.* [33] and Coles and Furrer [35] (see also Martens and de Muynck [29]). Both contributions formalize error and disturbance in terms of relative or conditional entropies, and derive their results from entropic uncertainty relations for state preparation which incorporate the effects of quantum entanglement [6, 44]. They differ in the choice of the entropic measure and the choice of the state on which the entropic terms are evaluated. Buscemi *et al.* find state-independent error-disturbance relations involving the von Neumann entropy, evaluated for input states which describe observable eigenstates chosen uniformly at random. As described in Sec. 6, the restriction to uniformly-random inputs is significant, and leads to a characterization of the average-case behavior of the device (averaged over the choice of input state), not the worst-case behavior as presented here. Meanwhile, Coles and Furrer make use of general Rényi-type entropies, hence also capturing the worst-case behavior. However, they are after a state-dependent error-disturbance relation which relates the amount of information a measurement device can extract from a state about the results of a *future* measurement of one observable to the amount of disturbance caused to other observable.

An important distinction between both these results and those presented here is the quantity appearing in the uncertainty bound, i.e. the quantification of complementarity of two observables. As both the aforementioned results are based on entropic state-preparation uncertainty relations, they each quantify complementarity by the largest overlap of the eigenstates of the two observables. This bound is trivial should the two

<sup>3</sup>Such an approach has been advocated by David Reeb (private communication).

observables share an eigenstate. However, a perfect joint measurement is clearly impossible even if the observables share all but two eigenvectors (if they share all but one, they necessarily share all eigenvectors). All three complementarity measures used here are nontrivial whenever not all eigenvectors are shared between the observables.

## 8 Conclusions

We have formulated simple, operational definitions of error and disturbance based on the probability of distinguishing the actual measurement apparatus from the relevant ideal apparatus by any testing procedure whatsoever. The resulting quantities are conceptually straightforward properties of the measurement apparatus, not any particular fixed experimental setup. We presented uncertainty relations for both joint measurability and the error-disturbance tradeoff, for both arbitrary finite-dimensional systems and for position and momentum. In the former case the bounds involve simple measures of the complementarity of two observables, while the latter involve the ratio of the desired position and momentum precisions  $\sigma_Q$  and  $\sigma_P$  to Planck's constant  $\hbar$ . We further showed that this operational approach has applications to quantum information processing and to wave-particle duality relations. Finally, we presented a detailed comparison of the relation of our results to previous work on uncertainty relations.

Several interesting questions remain open. One may inquire about the tightness of the bounds. The qubit example for conjugate observables discussed at the end of §3 shows that the finite-dimensional bounds of Theorem 2 are tight for small error  $\varepsilon_X$ , though no conclusion can be drawn from this example for small preparation disturbance. It would be interesting to check the tightness of the position and momentum bounds by computing the error and disturbance measures for a device described by a covariant measurement. For reasons of simplicity, we have not attempted to incorporate precision limits into the definitions of error and distinguishability of position and momentum. Doing so would lead to more conceptually satisfying bounds and perhaps remedy the fact that the measurement error-preparation disturbance bound is nontrivial even outside the Heisenberg limit. Bounds for other observables in infinite dimensions would also be quite interesting, for instance the mixed discrete/continuous case of energy and position of a harmonic oscillator. Restricting to covariant measurements, in finite or infinite dimensions, it would also be interesting to determine if entangled inputs improve the distinguishability measures, or whether calibration testing is sufficient. From the application in Corollary 1, it would appear that calibration is sufficient, but we have not settled the matter conclusively.

**Acknowledgements:** The authors are grateful to David Sutter, Paul Busch, Omar Fawzi, Fabian Furrer, Michael Walter and especially David Reeb and Reinhard Werner for helpful discussions. This work was supported by the Swiss National Science Foundation (through the NCCR ‘Quantum Science and Technology’ and grant no. 200020-135048) and the European Research Council (grant no. 258932). SH is funded by the German Excellence Initiative and the European Union Seventh Framework Programme under grant agreement no. 291763. He acknowledges additional support by DFG project no. K05430/1-1.

## References

- [1] W. Heisenberg, “Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik”, *Zeitschrift für Physik* **43**, 172–198 (1927) (page 1).
- [2] J. A. Wheeler and W. H. Zurek, *Quantum Theory and Measurement* (Princeton University Press, 1984) (page 1).
- [3] E. H. Kennard, “Zur Quantenmechanik einfacher Bewegungstypen”, *Zeitschrift für Physik* **44**, 326–352 (1927) (pages 1, 14).
- [4] H. P. Robertson, “The Uncertainty Principle”, *Physical Review* **34**, 163 (1929) (pages 1, 16).
- [5] H. Maassen and J. B. M. Uffink, “Generalized entropic uncertainty relations”, *Physical Review Letters* **60**, 1103 (1988) (page 1).
- [6] M. Berta, M. Christandl, R. Colbeck, J. M. Renes, and R. Renner, “The uncertainty principle in the presence of quantum memory”, *Nature Physics* **6**, 659–662 (2010), [arXiv:0909.0950 \[quant-ph\]](#) (pages 1, 3, 15, 17).
- [7] P. J. Coles, M. Berta, M. Tomamichel, and S. Wehner, “Entropic uncertainty relations and their applications”, *Reviews of Modern Physics* **89**, 015002 (2017), [arXiv:1511.04857 \[quant-ph\]](#) (pages 1, 3, 9).
- [8] E. Arthurs and J. L. Kelly, “On the Simultaneous Measurement of a Pair of Conjugate Observables”, *Bell System Technical Journal* **44**, 725–729 (1965) (page 1).
- [9] C. Y. She and H. Heffner, “Simultaneous Measurement of Noncommuting Observables”, *Physical Review* **152**, 1103–1110 (1966) (page 1).
- [10] E. B. Davies, *Quantum theory of open systems* (Academic Press, London, 1976) (pages 1, 3, 4).
- [11] S. T. Ali and E. Prugovečki, “Systems of imprimitivity and representations of quantum mechanics on fuzzy phase spaces”, *Journal of Mathematical Physics* **18**, 219–228 (1977) (page 1).

- [12] E. Prugovečki, “On fuzzy spin spaces”, *Journal of Physics A: Mathematical and General* **10**, 543 (1977) (page 1).
- [13] P. Busch, “Indeterminacy relations and simultaneous measurements in quantum theory”, *International Journal of Theoretical Physics* **24**, 63–92 (1985) (page 1).
- [14] P. Busch, “Unsharp reality and joint measurements for spin observables”, *Physical Review D* **33**, 2253–2261 (1986) (page 1).
- [15] E. Arthurs and M. S. Goodman, “Quantum Correlations: A Generalized Heisenberg Uncertainty Relation”, *Physical Review Letters* **60**, 2447–2449 (1988) (page 1).
- [16] H. Martens and W. M. de Muynck, “Towards a new uncertainty principle: quantum measurement noise”, *Physics Letters A* **157**, 441–448 (1991) (page 1).
- [17] S. Ishikawa, “Uncertainty relations in simultaneous measurements for arbitrary observables”, *Reports on Mathematical Physics* **29**, 257–273 (1991) (page 1).
- [18] M. G. Raymer, “Uncertainty principle for joint measurement of noncommuting variables”, *American Journal of Physics* **62**, 986–993 (1994) (page 1).
- [19] U. Leonhardt, B. Böhmer, and H. Paul, “Uncertainty relations for realistic joint measurements of position and momentum in quantum optics”, *Optics Communications* **119**, 296–300 (1995) (page 1).
- [20] D. M. Appleby, “Concept of Experimental Accuracy and Simultaneous Measurements of Position and Momentum”, *International Journal of Theoretical Physics* **37**, 1491–1509 (1998), [arXiv:quant-ph/9803046](#) (page 1).
- [21] M. J. W. Hall, “Prior information: How to circumvent the standard joint-measurement uncertainty relation”, *Physical Review A* **69**, 052113 (2004), [arXiv:quant-ph/0309091](#) (page 1).
- [22] R. F. Werner, “The uncertainty relation for joint measurement of position and momentum”, *Quantum Information and Computation* **4**, 546–562 (2004), [arXiv:quant-ph/0405184](#) (pages 1, 14, 17).
- [23] M. Ozawa, “Uncertainty relations for joint measurements of noncommuting observables”, *Physics Letters A* **320**, 367–374 (2004) (page 1).
- [24] Y. Watanabe, T. Sagawa, and M. Ueda, “Uncertainty relation revisited from quantum estimation theory”, *Physical Review A* **84**, 042121 (2011), [arXiv:1010.3571 \[quant-ph\]](#) (page 1).
- [25] P. Busch, P. Lahti, and R. F. Werner, “Proof of Heisenberg’s Error-Disturbance Relation”, *Physical Review Letters* **111**, 160405 (2013), [arXiv:1306.1565 \[quant-ph\]](#) (pages 1, 17).
- [26] P. Busch, P. Lahti, and R. F. Werner, “Heisenberg uncertainty for qubit measurements”, *Physical Review A* **89**, 012129 (2014), [arXiv:1311.0837 \[quant-ph\]](#) (pages 1, 17).
- [27] P. Busch, P. Lahti, and R. F. Werner, “Measurement uncertainty relations”, *Journal of Mathematical Physics* **55**, 042111 (2014), [arXiv:1312.4392 \[quant-ph\]](#) (pages 1, 17).
- [28] V. B. Braginsky and F. Y. Khalili, *Quantum Measurement* (Cambridge University Press, 1992) (page 1).
- [29] H. Martens and W. M. de Muynck, “Disturbance, conservation laws and the uncertainty principle”, *Journal of Physics A: Mathematical and General* **25**, 4887 (1992) (pages 1, 17).
- [30] M. Ozawa, “Universally valid reformulation of the Heisenberg uncertainty principle on noise and disturbance in measurement”, *Physical Review A* **67**, 042105 (2003), [arXiv:quant-ph/0207121](#) (page 1).
- [31] Y. Watanabe and M. Ueda, “Quantum Estimation Theory of Error and Disturbance in Quantum Measurement”, (2011), [arXiv:1106.2526 \[quant-ph\]](#) (page 1).
- [32] C. Branciard, “Error-tradeoff and error-disturbance relations for incompatible quantum measurements”, *Proceedings of the National Academy of Sciences* **110**, 6742–6747 (2013), [arXiv:1304.2071 \[quant-ph\]](#) (page 1).
- [33] F. Buscemi, M. J. W. Hall, M. Ozawa, and M. M. Wilde, “Noise and Disturbance in Quantum Measurements: An Information-Theoretic Approach”, *Physical Review Letters* **112**, 050401 (2014), [arXiv:1310.6603 \[quant-ph\]](#) (pages 1, 15, 17).
- [34] A. C. Ipsen, “Error-disturbance relations for finite dimensional systems”, (2013), [arXiv:1311.0259 \[math-ph\]](#) (pages 1, 17).
- [35] P. J. Coles and F. Furrer, “State-dependent approach to entropic measurement–disturbance relations”, *Physics Letters A* **379**, 105–112 (2015), [arXiv:1311.7637 \[quant-ph\]](#) (pages 1, 17).
- [36] M. Ozawa, “Uncertainty relations for noise and disturbance in generalized quantum measurements”, *Annals of Physics* **311**, 350–416 (2004) (page 1).
- [37] P. Busch, P. Lahti, and R. F. Werner, “Quantum root-mean-square error and measurement uncertainty relations”, *Reviews of Modern Physics* **86**, 1261–1281 (2014), [arXiv:1312.4393 \[quant-ph\]](#) (page 1).
- [38] D. M. Appleby, “Quantum Errors and Disturbances: Response to Busch, Lahti and Werner”, *Entropy* **18**, 174 (2016), [arXiv:1602.09002 \[quant-ph\]](#) (page 1).
- [39] M. Ozawa, “Disproving Heisenberg’s error-disturbance relation”, (2013), [arXiv:1308.3540 \[quant-ph\]](#) (page 1).
- [40] D. Kretschmann, D. Schlingemann, and R. Werner, “The Information-Disturbance Tradeoff and the Continuity of Stinespring’s Representation”, *IEEE Transactions on Information Theory* **54**, 1708–1717 (2008), [arXiv:quant-ph/0605009](#) (pages 2, 5, 16).

- [41] D. Kretschmann, D. Schlingemann, and R. F. Werner, “A continuity theorem for Stinespring’s dilation”, *Journal of Functional Analysis* **255**, 1889–1904 (2008), [arXiv:0710.2495 \[quant-ph\]](#) (pages 2, 5).
- [42] B.-G. Englert, “Fringe Visibility and Which-Way Information: An Inequality”, *Physical Review Letters* **77**, 2154 (1996) (pages 2, 16, 25).
- [43] J. M. Renes and J.-C. Boileau, “Conjectured Strong Complementary Information Tradeoff”, *Physical Review Letters* **103**, 020402 (2009), [arXiv:0806.3984 \[quant-ph\]](#) (page 3).
- [44] M. Tomamichel and R. Renner, “Uncertainty Relation for Smooth Entropies”, *Physical Review Letters* **106**, 110506 (2011), [arXiv:1009.2015 \[quant-ph\]](#) (pages 3, 15, 17).
- [45] M. Tomamichel, C. C. W. Lim, N. Gisin, and R. Renner, “Tight finite-key analysis for quantum cryptography”, *Nature Communications* **3**, 634 (2012), [arXiv:1103.4130 \[quant-ph\]](#) (page 3).
- [46] F. G. Lacerda, J. M. Renes, and R. Renner, “Classical leakage resilience from fault-tolerant quantum computation”, (2014), [arXiv:1404.7516 \[quant-ph\]](#) (pages 3, 15).
- [47] K. Kraus, *States, Effects, and Operations: Fundamental Notions of Quantum Theory*, Lecture notes in physics 190 (Springer-Verlag, Berlin, 1983) (page 3).
- [48] R. F. Werner, “Quantum Information Theory — an Invitation”, in *Quantum Information*, Springer Tracts in Modern Physics 173 (Springer Berlin Heidelberg, 2001), pp. 14–57, [arXiv:quant-ph/0101061](#) (page 3).
- [49] M. M. Wolf, “Quantum Channels and Operations: A Guided Tour”, (2012) (pages 3, 21).
- [50] C. Bény and F. Richter, “Algebraic approach to quantum theory: a finite-dimensional guide”, (2015), [arXiv:1505.03106 \[quant-ph\]](#) (page 3).
- [51] A. Y. Kitaev, “Quantum computations: Algorithms and error correction”, *Russian Mathematical Surveys* **52**, 1191–1249 (1997) (pages 4, 17).
- [52] V. Paulsen, *Completely Bounded Maps and Operator Algebras*, Vol. 78, Cambridge Studies in Advanced Mathematics (Cambridge University Press, Jan. 15, 2003), 320 pp. (page 4).
- [53] A. Gilchrist, N. K. Langford, and M. A. Nielsen, “Distance measures to compare real and ideal quantum processes”, *Physical Review A* **71**, 062310 (2005), [arXiv:quant-ph/0408063](#) (page 4).
- [54] J. Watrous, “Semidefinite Programs for Completely Bounded Norms”, *Theory of Computing* **5**, 217–238 (2009), [arXiv:0901.4709 \[quant-ph\]](#) (pages 4, 21).
- [55] J. Watrous, “Simpler semidefinite programs for completely bounded norms”, *Chicago Journal of Theoretical Computer Science* **2013**, 8 (2013), [arXiv:1207.5726 \[quant-ph\]](#) (page 4).
- [56] W. F. Stinespring, “Positive functions on  $C^*$ -algebras”, *Proceedings of the American Mathematical Society* **6**, 211–216 (1955) (page 4).
- [57] P. J. Coles and M. Piani, “Improved entropic uncertainty relations and information exclusion relations”, *Physical Review A* **89**, 022112 (2014), [arXiv:1307.4265 \[quant-ph\]](#) (page 9).
- [58] P. W. Shor and J. Preskill, “Simple Proof of Security of the BB84 Quantum Key Distribution Protocol”, *Physical Review Letters* **85**, 441 (2000), [arXiv:quant-ph/0003004](#) (page 15).
- [59] I. Devetak, “The private classical capacity and quantum capacity of a quantum channel”, *IEEE Transactions on Information Theory* **51**, 44–55 (2005), [arXiv:quant-ph/0304127](#) (page 15).
- [60] J. M. Renes, “Duality of privacy amplification against quantum adversaries and data compression with quantum side information”, *Proceedings of the Royal Society A* **467**, 1604–1623 (2011), [arXiv:1003.0703 \[quant-ph\]](#) (page 15).
- [61] J. M. Renes, “The Physics of Quantum Information: Complementarity, Uncertainty, and Entanglement”, *Habilitation* (TU Darmstadt, 2012), [arXiv:1212.2379 \[quant-ph\]](#) (page 15).
- [62] J. M. Renes, “Uncertainty relations and approximate quantum error correction”, *Physical Review A* **94**, 032314 (2016), [arXiv:1605.01420 \[quant-ph\]](#) (page 16).
- [63] P. J. Coles, J. Kaniewski, and S. Wehner, “Equivalence of wave-particle duality to entropic uncertainty”, *Nature Communications* **5**, 5814 (2014), [arXiv:1403.4687 \[quant-ph\]](#) (page 16).
- [64] P. J. Coles, “Entropic framework for wave-particle duality in multipath interferometers”, *Physical Review A* **93**, 062111 (2016), [arXiv:1512.09081 \[quant-ph\]](#) (page 16).
- [65] K. Korzekwa, D. Jennings, and T. Rudolph, “Operational constraints on state-dependent formulations of quantum error-disturbance trade-off relations”, *Physical Review A* **89**, 052108 (2014), [arXiv:1311.5506 \[quant-ph\]](#) (page 16).
- [66] A. Barchielli, M. Gregoratti, and A. Toigo, “Measurement uncertainty relations for discrete observables: Relative entropy formulation”, (2016), [arXiv:1608.01986 \[math-ph\]](#) (page 17).
- [67] M. F. Sacchi, “Entanglement can enhance the distinguishability of entanglement-breaking channels”, *Physical Review A* **72**, 014305 (2005), [arXiv:quant-ph/0505174](#) (page 17).
- [68] V. P. Belavkin, “Optimal multiple quantum statistical hypothesis testing”, *Stochastics* **1**, 315 (1975) (page 24).
- [69] P. Hausladen and W. K. Wootters, “A ‘Pretty Good’ Measurement for Distinguishing Quantum States”, *Journal of Modern Optics* **41**, 2385 (1994) (page 24).

## A Entanglement improves the distinguishability of measurements

Here we give an example of two measurements whose distinguishability is improved by entanglement. Let  $\mathcal{E}_1$  be a measurement in an arbitrary chosen basis  $|b_0\rangle, |b_1\rangle$ , and  $|b_2\rangle$ , and define  $\mathcal{E}_2$  be measurement in the basis given by  $|\theta_0\rangle = \frac{1}{3}(2|b_0\rangle + 2|b_1\rangle - |b_2\rangle)$ ,  $|\theta_1\rangle = \frac{1}{3}(-1|b_0\rangle + 2|b_1\rangle + 2|b_2\rangle)$  and  $|\theta_2\rangle = \frac{1}{3}(2|b_0\rangle - |b_1\rangle + 2|b_2\rangle)$ . Using  $T_k = |b_k\rangle\langle b_k| - |\theta_k\rangle\langle\theta_k|$ , the largest distinguishability to be had without entanglement is given by

$$\delta'(\mathcal{E}_1, \mathcal{E}_2) = \max_{\varrho} \frac{1}{2} \sum_{k=0}^2 |\text{Tr}[\varrho T_k]| \quad (52a)$$

$$= \max_{\varrho} \max_{\{s_k=\pm 1\}} \frac{1}{2} \sum_{k=0}^2 \text{Tr}[s_k T_k \varrho] \quad (52b)$$

$$= \max_{\{s_k=\pm 1\}} \left\| \sum_{k=0}^2 s_k T_k \right\|_{\infty}. \quad (52c)$$

Checking the eight combinations of  $s_k$ , one easily finds that the maximum value is  $\sqrt{5}/3$ .

Meanwhile, if we use the state

$$\varrho = \frac{1}{6} \begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix} \quad (53)$$

to define  $\Psi = (\mathbb{1} \otimes \sqrt{\varrho})\Omega(\mathbb{1} \otimes \sqrt{\varrho})$  for  $\Omega$  the projector onto  $|\Omega\rangle = \sum_k |b_k\rangle \otimes |b_k\rangle$ , then

$$\delta(\mathcal{E}_1, \mathcal{E}_2) \geq \frac{1}{2} \sum_{k=0}^2 \left\| \text{Tr}_1[(T_k \otimes \mathbb{1})\Psi] \right\|_1. \quad (54)$$

Direct calculation shows that  $\delta(\mathcal{E}_1, \mathcal{E}_2) \geq \sqrt{3}/2$ . Thus, there exist projective measurements for which  $\delta(\mathcal{E}_1, \mathcal{E}_2) > \delta'(\mathcal{E}_1, \mathcal{E}_2)$ .

## B Computing error and disturbance by convex optimization

Here we detail how to compute the error and disturbance quantities via semidefinite programming and calculate these for the nonideal qubit  $X$  measurement example. Given a Hilbert space  $\mathcal{H}$  with basis  $\{|k\rangle\}_{k=1}^d$ , define, just as above,  $|\Omega\rangle = \sum_{k=1}^d |k\rangle \otimes |k\rangle \in \mathcal{H} \otimes \mathcal{H}$ . Then, for any channel  $\mathcal{E}$ , let  $\mathbb{C}$  denote the Choi mapping of  $\mathcal{E}^*$  to an unnormalized bipartite state,

$$\mathbb{C}(\mathcal{E}) := \mathcal{E}^* \otimes \mathcal{I}(|\Omega\rangle\langle\Omega|) \in \mathcal{B}(\mathcal{H}_B \otimes \mathcal{H}_A). \quad (55)$$

The action of the channel can be compactly expressed in terms of the Choi operator as  $\mathcal{E}_{A \rightarrow B}(\Lambda_B) = \text{Tr}_B[\Lambda_B \mathbb{C}(\mathcal{E})_{BA}]$  or in the Schrödinger picture as  $\mathcal{E}_{A \rightarrow B}^*(\varrho_A) = \text{Tr}_A[\mathbb{C}(\mathcal{E})_{BA} \varrho_A^T]$ , where the transpose is taken in the basis defining  $\mathbb{C}$  (see, e.g. [49]). The cb norm can then be expressed in primal and dual form as [54]

$$\frac{1}{2} \|\mathcal{E}_{A \rightarrow B}\|_{\text{cb}} = \underset{K, \varrho}{\text{maximum}} \quad \text{Tr}[\mathbb{C}(\mathcal{E})_{BA} K_{BA}] \quad (56)$$

$$\text{subject to} \quad K_{BA} - \mathbb{1}_B \otimes \varrho_A \leq 0, \quad \text{Tr}[\varrho_A] \leq 1, \\ \varrho_A, K_{BA} \geq 0,$$

$$= \underset{T, \lambda}{\text{minimum}} \quad \lambda \quad (57)$$

$$\text{subject to} \quad T_{BA} \geq \mathbb{C}(\mathcal{E})_{BA}, \quad \lambda \mathbb{1}_A - T_A \geq 0, \\ T_{BA}, \lambda \geq 0.$$

Note that in the dual formulation the objective function is just the operator norm  $\|T_A\|_{\infty}$ . For infinite-dimensional systems the Choi operator does not have such a nice form, though it might be possible to formulate the cb norm of Gaussian channels as a tractable optimization.

The additional optimizations involving  $\mathcal{R}$  in the measures of error and disturbance are immediately compatible with the dual formulation in (57), and so these quantities can be cast as semidefinite programs. To

start, consider the error in measuring  $X$ . With  $Q_{XA} = C(Q_X)$  and  $E_{YBA} = C(\mathcal{E}_{A \rightarrow YB})$ , we have

$$\begin{aligned} \varepsilon_X(\mathcal{E}_{A \rightarrow YB}) &= \underset{T, \lambda, R}{\text{minimum}} \quad \lambda & (58) \\ \text{subject to} \quad & T_{XA} + \text{Tr}_Y[R_{XY}E_{YA}] \geq Q_{XA}, \quad \lambda \mathbb{1}_A - T_A \geq 0, \quad R_Y = \mathbb{1}_Y, \\ & \lambda, T_{XA}, R_{XY} \geq 0. \end{aligned}$$

Without loss of generality, we may restrict the operator  $T_{XA}$  to be a hybrid classical-quantum operator, classical on  $X$ , and of course  $R_{XY}$  is classical on both systems. This is also the reason it is unnecessary to transpose  $Y$  in  $\text{Tr}_Y[R_{XY}E_{YA}]$ . Further symmetries of  $Q_{XA}$  and  $E_{XA}$  can be quite helpful in simplifying the program, but we will not pursue this further here. The associated primal form is as follows.

$$\begin{aligned} \varepsilon_X(\mathcal{E}_{A \rightarrow YB}) &= \underset{K, \varrho, L}{\text{maximum}} \quad \text{Tr}[Q_{XA}K_{XA}] - \text{Tr}[L_Y] & (59) \\ \text{subject to} \quad & K_{XA} - \mathbb{1}_X \otimes \varrho_A \leq 0, \quad \text{Tr}[\varrho_A] \leq 1, \quad \text{Tr}_A[E_{YA}K_{XA}] - L_Y \otimes \mathbb{1}_X \leq 0, \\ & \varrho_A, K_{XA} \geq 0, L_Y = L_Y^*. \end{aligned}$$

In writing an equality we have assumed that the duality gap is zero. But this is easy enough to show using the Slater condition, namely by ensuring that the value of the minimization is finite and that there exists a strictly feasible set of maximization variables. The former holds because  $\varepsilon_X$  is the infimum of the distinguishability, and hence  $\varepsilon_X(\mathcal{E}) \geq 0$ . Meanwhile, a strictly feasible set of variables in (59) is given by  $K = \frac{1}{2}k\mathbb{1}$ ,  $\varrho = k\mathbb{1}$ , and  $L = kE_Y$  for  $k < 1/\dim(A)$ .

To formulate the measurement disturbance  $\nu_Z(\mathcal{E}_{A \rightarrow YB})$  we are interested in  $C(\mathcal{E}\mathcal{R}\mathcal{T}_Y Q_Z)$ , which can be expressed as a linear map on  $R_{ABY}$ :

$$C(\mathcal{E}\mathcal{R}\mathcal{T}_Y Q_Z) = \text{Tr}_{A'YB}[Q_{ZA'}R_{A'YB}^{T_{A'}}E_{YBA}^{T_B}] \quad (60a)$$

$$= \text{Tr}_{A'YB}[R_{A'YB}Q_{ZA'}^{T_{A'}}E_{YBA}^{T_B}]. \quad (60b)$$

In the second step we have transposed the  $A'$  system in the first. Then we have

$$\nu_Z(\mathcal{E}_{A \rightarrow YB}) = \underset{T, \lambda, R}{\text{minimum}} \quad \lambda \quad (61)$$

$$\begin{aligned} \text{subject to} \quad & T_{ZA} + \text{Tr}_{A'YB}[R_{A'YB}Q_{ZA'}^{T_{A'}}E_{YBA}^{T_B}] \geq Q_{ZA}, \quad \lambda \mathbb{1}_A - T_A \geq 0, \quad R_{YB} = \mathbb{1}_{YB}, \\ & \lambda, T_{ZA}, R_{A'YB} \geq 0, \end{aligned}$$

$$= \underset{K, \varrho, L}{\text{maximum}} \quad \text{Tr}[Q_{ZA}K_{ZA}] - \text{Tr}[L_{YB}] \quad (62)$$

$$\begin{aligned} \text{subject to} \quad & K_{ZA} - \mathbb{1}_Z \otimes \varrho_A \leq 0, \quad \text{Tr}[\varrho_A] \leq 1, \quad \text{Tr}_{ZA}[Q_{ZA'}E_{YBA}K_{ZA}] - \mathbb{1}_A \otimes L_{YB} \leq 0, \\ & \varrho_A, K_{ZA} \geq 0, L_{YB} = L_{YB}^*. \end{aligned}$$

Here we have absorbed the transposes over  $A'$  and  $B$  into  $\mathbb{1}_{A'}$  and the definition of  $L_{YB}$ , since this does not affect Hermiticity or the value of the objective function. Strong duality is essentially the same as before: The minimization is finite and we can choose  $K = \frac{1}{2}k\mathbb{1}$  and  $\varrho = k\mathbb{1}$ . Then in the third constraint we have  $\text{Tr}_{ZA}[Q_{ZA'}E_{YBA}K_{ZA}] = \frac{1}{2}k\mathbb{1}_{A'} \otimes E_{YB}$  since  $Q_Z$  is unital. Setting  $L = kE_{YB}$  gives a strictly feasible set.

Finally, we come to the two preparation disturbance measures. The first is simply

$$\eta_Z(\mathcal{E}_{A \rightarrow YB}) = \underset{T, \lambda, R}{\text{minimum}} \quad \lambda \quad (63)$$

$$\begin{aligned} \text{subject to} \quad & T_{AZ} + \text{Tr}_{YB A'}[R_{AYB}E_{YBA}^{T_B}P_{A'Z}^{T_{A'}}] \geq P_{AZ}, \quad \lambda \mathbb{1}_Z - T_Z \geq 0, \quad R_{YB} = \mathbb{1}_{YB}, \\ & \lambda, T_{AZ}, R_{AYB} \geq 0, \end{aligned}$$

$$= \underset{K, \varrho, L}{\text{maximum}} \quad \text{Tr}[P_{AZ}K_{AZ}] - \text{Tr}[L_{YB}] \quad (64)$$

$$\begin{aligned} \text{subject to} \quad & K_{AZ} - \mathbb{1}_A \otimes \varrho_Z \leq 0, \quad \text{Tr}[\varrho_Z] \leq 1, \quad \text{Tr}_{AZ}[E_{YBA}P_{A'Z}^{T_{A'}}K_{AZ}] - \mathbb{1}_A \otimes L_{YB} \leq 0, \\ & \varrho_Z, K_{AZ} \geq 0, L_{YB} = L_{YB}^*. \end{aligned}$$

Here we have absorbed the transpose on  $B$  into the definition of  $L_{YB}$  since this doesn't affect Hermiticity or the value of the objective function. Strong duality holds as before, and also for the demerit measure which

reads

$$\frac{d-1}{d} - \widehat{\eta}_Z(\mathcal{E}_{A \rightarrow YB}) = \underset{T, \lambda, \sigma}{\text{minimum}} \quad \lambda \quad (65)$$

$$\text{subject to} \quad T_{YBZ} + \sigma_{YB} \otimes \mathbb{1}_Z \geq \text{Tr}_A[E_{YBA} P_{AZ}^{T_A}], \quad \lambda \mathbb{1}_Z - T_Z \geq 0, \quad \text{Tr}[\sigma_{YB}] = 1, \\ \lambda, T_{YBZ}, \sigma_{YB} \geq 0,$$

$$= \underset{K, \varrho, \mu}{\text{maximum}} \quad \text{Tr}[E_{YBA} P_{AZ}^{T_A} K_{YBZ}] - \mu \quad (66)$$

$$\text{subject to} \quad K_{YBZ} - \mathbb{1}_{YB} \otimes \varrho_Z \leq 0, \quad \text{Tr}[\varrho_Z] \leq 1, \quad K_{YB} - \mu \mathbb{1}_{YB} \leq 0, \\ \varrho_Z, K_{YBZ} \geq 0, \mu \in \mathbb{R}.$$

Now let us consider the particular example described in the main text, a suboptimal  $X$  measurement. Suppose we use  $|\varphi_x\rangle$  from the ideal  $X$  measurement to define the Choi operator. After a bit of calculation, one finds that the Choi operator  $E_{YBA}$  of  $\mathcal{E}_{YB|A}$  is given by

$$E_{YBA} = |b_0\rangle\langle b_0|_Y \otimes |\Psi\rangle\langle\Psi|_{BA} + |b_1\rangle\langle b_1|_Y \otimes (\sigma_z \otimes \sigma_z)|\Psi\rangle\langle\Psi|_{BA} (\sigma_z \otimes \sigma_z), \quad (67)$$

where  $|\Psi\rangle = \cos\frac{\theta}{2}|\varphi_0\rangle \otimes |\varphi_0\rangle + \sin\frac{\theta}{2}|\varphi_1\rangle \otimes |\varphi_1\rangle$ . Tracing out  $B$  gives the Choi operator of just the measurement result  $Y$ ,  $E_{YA} = \sum_x |b_x\rangle\langle b_x|_Y \otimes \Lambda_x$ , with  $\Lambda_x = \frac{1}{2}\mathbb{1} + \frac{1}{2}(-1)^x \cos\theta \sigma_x$ .

To compute the measurement error  $\varepsilon_X(\mathcal{E})$ , suppose that no recovery operation is applied, i.e. the outcome  $Y$  is treated as  $X$ . Then we can work with  $E_{XA}$  and dispense with  $R$  so that the third constraint in (58) is satisfied. To satisfy the first constraint, choose  $T_{XA}$  to be the positive part of  $Q_{XA} - E_{XA}$ . This gives  $T_{XA} = \frac{1}{2}(1 - \cos\theta) \sum_x |b_x\rangle\langle b_x| \otimes |\varphi_x\rangle\langle\varphi_x|$ ; consequently,  $T_A = \frac{1}{2}(1 - \cos\theta)\mathbb{1}_A$  and therefore  $\varepsilon_X(\mathcal{E}) \leq \frac{1}{2}(1 - \cos\theta)$ . On the other hand,  $K_{XA} = \frac{1}{2}Q_{XA}$  and  $\varrho_A = \frac{1}{2}\mathbb{1}_A$  satisfy the first two constraints in (59). The last constraint involves the quantity  $\text{Tr}_A[E_{YA} K_{XA}] = \frac{1}{4} \sum_{x,y} |b_y\rangle\langle b_y|_Y \otimes |b_x\rangle\langle b_x|_X (1 + (-1)^{x+y} \cos\theta)$  and can therefore be satisfied by choosing  $L_Y = \frac{1}{4}(1 + \cos\theta)\mathbb{1}_Y$ . Evaluating the objective function gives  $\varepsilon_X(\mathcal{E}) \geq \frac{1}{2}(1 - \cos\theta)$ .

Note that the choice of  $K_{XA}$  corresponds to the unentangled test of randomly inputting  $|\varphi_x\rangle$  and checking that the result is  $x$ . We could have anticipated that unentangled tests would be sufficient in this case, since the optimal and actual measurements are both diagonal in the  $\sigma_x$  basis: Any input state can be freely dephased in this basis, thus removing any entanglement.

Next, consider the measurement disturbance  $\nu_Z(\mathcal{E})$ . Proceeding as with measurement error, suppose that no recovery operation is applied, so that the output  $B$  is just regarded as  $A' \simeq A$  and the third constraint in (61) is trivially satisfied. For the first constraint we need only the operator  $\text{Tr}_{Y A'}[Q_{Z A'} E_{Y A' A}^{T_{A'}}]$ , and after some calculation we find that it equals  $\sum_z |b_z\rangle\langle b_z| \otimes \Gamma_z$  with  $\Gamma_z = \frac{1}{2}(\mathbb{1} + (-1)^z \sin\theta \sigma_z)$ . Thus, the optimization is just like that of  $\varepsilon_X(\mathcal{E})$ , but with  $\cos\theta$  replaced by  $\sin\theta$ . Hence  $\nu_Z(\mathcal{E}) \leq \frac{1}{2}(1 - \sin\theta)$ . To show the other inequality from the maximization form (62) also proceeds as before, starting with  $K_{ZA} = \frac{1}{2}Q_{ZA}$  and  $\varrho_A = \frac{1}{2}\mathbb{1}_A$ . For the third constraint a bit of calculation shows

$$\text{Tr}_{ZA}[Q_{ZA} E_{YBA} K_{ZA}] = \frac{1}{4} \sum_{x,z} |b_z\rangle\langle b_z|_{A'} \otimes |b_x\rangle\langle b_x|_Y \otimes (\sigma_x^z \sigma_z^x |\psi\rangle\langle\psi| \sigma_z^x \sigma_x^z)_B, \quad (68)$$

with  $|\psi\rangle = \frac{1}{\sqrt{2}}(\sqrt{1 + \sin\theta}|\theta_0\rangle + \sqrt{1 - \sin\theta}|\theta_1\rangle)$ . Choosing

$$L_{YB} = \frac{1}{8} \sum_x |b_x\rangle\langle b_x|_Y \otimes ((1 + \sin\theta)\mathbb{1} + (-1)^x \cos\theta \sigma_x)_B \quad (69)$$

satisfies the constraints, and the objective function becomes  $\frac{1}{2}(1 - \sin\theta)$ . As with  $\varepsilon_X(\mathcal{E})$ , entangled inputs do not increase the distinguishability in this particular case.

A trivial recovery map also optimizes  $\eta_Z(\mathcal{E})$ . To see this, set  $K_{AZ} = \frac{1}{2}P_{AZ}$  and  $\varrho_Z = \frac{1}{2}\mathbb{1}_Z$ . Then in the third constraint of (64) we have  $\text{Tr}_{A'}[E_{YBA} P_{AZ}^{T_{A'}} K_{AZ}]$ , which is precisely the same as (68) with  $A'$  replaced by  $A$ . Hence, if we choose  $L_{YB}$  as in (69), we obtain the lower bound  $\eta_Z(\mathcal{E}) \geq \frac{1}{2}(1 - \sin\theta)$ . To establish optimality, suppose  $\mathcal{R}$  does nothing but discard the  $Y$  system. In the minimization (63) we then have  $\text{Tr}_{Y A'}[E_{Y A' A}^{T_{A'}}]$ , which is the same as  $\text{Tr}_{Y A'}[Q_{Z A'} E_{Y A' A}^{T_{A'}}]$  from  $\nu_Z(\mathcal{E})$ . Proceeding as there, we find the matching upper bound.

Finally, consider  $\widehat{\eta}_Z(\mathcal{E})$ . Here there are two possible outputs of  $\mathcal{P}_Z \mathcal{E}$ , call them  $\xi_0$  and  $\xi_1$ . It is not difficult to show that for arbitrary  $\xi_z$  the distinguishability is precisely  $\widehat{\eta}_Z(\mathcal{E}) = \frac{1}{2}(1 - \delta(\xi_0, \xi_1))$ . On the one hand, we can simply pick the output of  $\mathcal{C}$  to be  $\xi = \frac{1}{2}(\xi_0 + \xi_1)$ . Then, with  $T$  in (65) the positive part of  $\sum_z |z\rangle\langle z|_Z \otimes (\xi_z - \xi)$ , the objective function becomes  $\frac{1}{2}\delta(\xi_0, \xi_1)$ . On the other hand, in (66) we can choose

$K_{YBZ} = \frac{1}{2}|0\rangle\langle 0|_Z \otimes \Lambda_{YB} + \frac{1}{2}|1\rangle\langle 1|_Z \otimes (\mathbb{1} - \Lambda)_{YB}$ , for  $\Lambda$  the projector onto the nonnegative part of  $\xi_0 - \xi_1$ . Then  $\mu = \frac{1}{2}$  and  $\varrho = \frac{1}{2}\mathbb{1}$  are feasible and lead again to the same objective function. In this particular case the two states are  $\xi_1 = \sigma_z \xi_0 \sigma_z$  and  $\xi_0 = \frac{1}{2} \sum_x |b_x\rangle\langle b_x|_Y \otimes \sigma_x |\psi\rangle\langle \psi|_{\sigma_x}$ , which yields  $\delta(\xi_0, \xi_1) = \sin \theta$  and hence  $\hat{\eta}_Z(\mathcal{E}) = \frac{1}{2}(1 - \sin \theta)$ .

### C Counterexample channel

Here we present the calculations involved in §6.1 Let  $|\xi_z\rangle_{BD} = \sum_y \sqrt{p_{yz}} |y\rangle_B |y\rangle_D$ . Then the isometry is just

$$V = \sum_z |\xi_z\rangle_{|z\rangle_C} \langle z|_A. \quad (70)$$

Observe that the action on  $|\tilde{x}\rangle$  states leads to symmetric output in  $BC$ :

$$V |\tilde{x}\rangle = \frac{1}{\sqrt{d}} \sum_z \omega^{xz} V |z\rangle \quad (71a)$$

$$= \frac{1}{\sqrt{d}} \sum_z \omega^{xz} |\xi_z\rangle_{BD} |z\rangle_C \quad (71b)$$

$$= Z_C^x \frac{1}{\sqrt{d}} \sum_z |\xi_z\rangle_{BD} |z\rangle_C. \quad (71c)$$

Therefore, the probability of incorrectly identifying any particular input state is the same as any other, and we can consider the case that the input  $x$  value is chosen uniformly at random. We can further simplify the  $BC$  output by defining  $p_y = \frac{1}{d} \sum_z p_{yz}$  and

$$|\eta_y\rangle = \frac{1}{\sqrt{d}} \sum_z \sqrt{p_{yz}/p_y} |z\rangle, \quad (72)$$

which is a normalized state on  $\mathcal{H}_C$  for each  $y$ . Then we have

$$V |\tilde{x}\rangle = Z_C^x \sum_y \sqrt{p_y} |\eta_y\rangle_C |y\rangle_B |y\rangle_D. \quad (73)$$

Ignoring the  $D$  system will produce a classical-quantum state, with system  $B$  recording the classical value  $y$ , which occurs with probability  $p_y$ , and  $C$  the quantum state  $Z^x |\eta_y\rangle$ . The optimal measurement therefore has elements  $\Lambda_x$  of the form  $\Lambda_x = \sum_y |y\rangle\langle y|_B \otimes (\Gamma_{x,y})_C$  for some set of POVMs  $\{\Gamma_{x,y}\}_y$ . In every sector of fixed  $y$  value, the measurement has to distinguish between a set of pure states occurring with equal probabilities. Therefore, by a result going back to Belavkin, the optimal measurement is the so-called ‘‘pretty good measurement’’ [68, 69]. This has measurement elements  $\Gamma_{x,y}$  which project onto the orthonormal states  $|\mu_{x,y}\rangle = S^{-1/2} Z^x |\eta_y\rangle$ , where  $S = \sum_x Z^x |\eta_y\rangle\langle \eta_y| Z^{-x}$ . It is easy to work out that  $S = \sum_x (p_{yz}/p_y) |z\rangle\langle z|$ , and thus  $|\mu_{x,y}\rangle = |\tilde{x}\rangle$  for all  $y$ . Hence, we can in fact dispense with the  $B$  system altogether, since the particular value of  $y$  does not alter the optimal measurement. The average guessing probability is thus

$$p_{\text{guess}} = \frac{1}{d} \sum_{x,y} p_y |\langle \tilde{x} | Z^x |\eta_y\rangle|^2 \quad (74a)$$

$$= \sum_y p_y |\langle \tilde{0} | \eta_y\rangle|^2 \quad (74b)$$

$$= \frac{1}{d^2} \sum_y \left( \sum_z \sqrt{p_{yz}} \right)^2, \quad (74c)$$

as intended.

### D Englert’s complementarity relation

Here we describe Englert’s setup in our formalism and establish (49). He considers a Mach-Zehnder interferometer with a relative phase shift between the two arms and additional which-way detectors in each arm. To the two possible paths inside the interferometer we may associate the (orthogonal) eigenstates  $|\vartheta_z\rangle$  of an observable  $Z$ , with  $z \in \{0, 1\}$ . For simplicity, we assume  $Z$  has eigenvalues  $(-1)^z$ . The action of a relative  $\varphi$  phase shift is described by the unitary  $U_{\text{PS}} = \sum_{z=0}^1 e^{iz\varphi} |\vartheta_z\rangle\langle \vartheta_z|$ . It will prove convenient to choose  $\varphi = 0$  below, but we leave it arbitrary for now. Meanwhile, the which-way detectors can be described as producing

different states of an ancilla system, depending on which path the photon takes. For pure ancilla states  $|\gamma_z\rangle$ , the detector corresponds to the isometry  $U_{\text{WW}} = \sum_{z=0}^1 |\vartheta_z\rangle\langle\vartheta_z|_Q \otimes |\gamma_z\rangle_A$ , where  $A$  denotes the ancilla and  $Q$  the system itself, which Englert terms a “quanton”.

Ignoring the phase shifts associated with reflection, the output modes of a symmetric (50/50) beamsplitter are related to the input modes by the unitary  $U_{\text{BS}} = \sum_{z=0}^1 |\vartheta_z\rangle\langle\varphi_z|$ , with  $|\varphi_x\rangle = \frac{1}{\sqrt{2}} \sum_z (-1)^{xz} |\vartheta_z\rangle$  for  $x \in \{0, 1\}$ . We may associate these states with the observable  $X$ , also taking eigenvalues  $(-1)^x$ . Observe that all three complementarity measures are  $\frac{1}{2}$ . The entire Mach-Zehnder device can be described by the isometry

$$U_{\text{MZ}} = U_{\text{BS}} U_{\text{PS}} U_{\text{WW}} U_{\text{BS}} \quad (75a)$$

$$= \sum_{x,z=0}^1 e^{ix\varphi} |\vartheta_z\rangle\langle\varphi_z| \langle\varphi_x|_Q \otimes |\gamma_x\rangle_A \quad (75b)$$

$$= \sum_{x=0}^1 e^{ix\varphi} |\varphi_x\rangle\langle\varphi_x|_Q \otimes |\gamma_x\rangle_A. \quad (75c)$$

When the ancilla is subsequently measured so as to extract information about the path, we may regard the whole operation as an apparatus  $\mathcal{E}_{\text{MZ}}$  with one quantum and one classical output.

The available “which-way” information, associated with particle-like behavior of  $Q$ , is characterized by the distinguishability  $D := \delta(\gamma_0, \gamma_1)$ . Given the particular form of  $U$  in (75), we may set  $\sin \theta = \langle\gamma_0|\gamma_1\rangle$  for  $\theta \in \mathbb{R}$  without loss of generality;  $D$  is then  $\cos \theta$ . This amounts to defining  $|\gamma_k\rangle = \cos \frac{\theta}{2} |k\rangle + \sin \frac{\theta}{2} |k+1\rangle$ , where the states  $\{|k\rangle\}_{k=0}^1$  form an orthonormal basis and arithmetic inside the ket is modulo two. Thus,  $\mathcal{E}_{\text{MZ}}$  with  $\varphi = 0$  is precisely the nonideal qubit  $X$  measurement  $\mathcal{E}$  considered in §3. We shall see momentarily that  $\varphi = 0$  can be chosen without loss of generality. Using (14) we have  $\varepsilon_X(\mathcal{E}_{\text{MZ}}) = \frac{1}{2}(1 - D)$  as claimed.

Meanwhile, the fringe visibility  $V$  is defined as the difference in probability (or population) in the two output modes of the interferometer, maximized over the choice of input state. Since  $Z = |\vartheta_0\rangle\langle\vartheta_0| - |\vartheta_1\rangle\langle\vartheta_1|$ , this is just

$$V = \max_{\rho} \left| \text{Tr}[(Z_Q \otimes \mathbb{1}_A) U_{\text{MZ}} \rho U_{\text{MZ}}^*] \right|. \quad (76)$$

A straightforward calculation yields  $U_{\text{MZ}}^*(Z_Q \otimes \mathbb{1}_A) U_{\text{MZ}} = \sin \theta (\cos \varphi Z + i \sin \varphi XZ)$ . It can be verified that  $(\cos \varphi Z + i \sin \varphi XZ)$  has eigenvalues  $\pm 1$ , and therefore  $V = \sin \theta$ . Thus,  $V^2 + D^2 = 1$  in this case (cf. [42, Eq. 11]). Note that  $\varphi$  does not appear in the visibility itself, justifying our choice of  $\varphi = 0$  above. By (15),  $\nu_Z(\mathcal{E}_{\text{MZ}}) = \eta_Z(\mathcal{E}_{\text{MZ}}) = \hat{\eta}_Z(\mathcal{E}_{\text{MZ}}) = \frac{1}{2}(1 - V)$ .