

Measurement disturbance and conservation laws in quantum mechanics

M. Hamed Mohammady^{1,2}, Takayuki Miyadera³, and Leon Loveridge⁴

¹QuIC, École Polytechnique de Bruxelles, CP 165/59, Université Libre de Bruxelles, 1050 Brussels, Belgium

²RCQI, Institute of Physics, Slovak Academy of Sciences, Dúbravská cesta 9, Bratislava 84511, Slovakia

³Department of Nuclear Engineering, Kyoto University, Nishikyo-ku, Kyoto 615-8540, Japan

⁴Quantum Technology Group, Department of Science and Industry Systems, University of South-Eastern Norway, 3616 Kongsberg, Norway

Measurement error and disturbance, in the presence of conservation laws, are analysed in general operational terms. We provide novel quantitative bounds demonstrating necessary conditions under which accurate or non-disturbing measurements can be achieved, highlighting an interesting interplay between incompatibility, unsharpness, and coherence. From here we obtain a substantial generalisation of the Wigner-Araki-Yanase (WAY) theorem. Our findings are further refined through the analysis of the fixed-point set of the measurement channel, some extra structure of which is characterised here for the first time.

1 Introduction

That measurements generally disturb quantum systems is one of the fundamental aspects of quantum mechanics. The consequences of this effect range from the foundational to the applied, sometimes entering in the guise of measurement “back-action”, playing a key role in quantum metrology, computation, and information processing [1–20]. Measurement disturbance can be seen when two observables are measured in succession, and the statistics of the second measurement depend on the first. While a well-known necessary condition for non-disturbance is that the pair of observables must be *compatible* [21, 22], further necessary conditions arise when the first measurement obeys a conservation law, i.e., when the interaction between the measured system and measuring apparatus conserves some total quantity such as energy, charge, or angular momentum. Indeed, the Wigner-Araki-Yanase (WAY) theorem states that when a single sharp observable is measured in succession, the first measurement will not disturb the second only if the measured observable commutes with the system part of a conserved quantity [23–25]. The same constraint holds for perfectly accurate measurements, and independently of disturbance, if the pointer observable of the apparatus obeys the “Yanase condition”, i.e., if it commutes with the apparatus part of a conserved quantity [26].

The WAY theorem has evolved over the years and continues to inspire research in a variety of directions (some recent examples are [27–34]), having impact also in other fields of research: for instance in quantum computing [35–37], the resource theories of asymmetry [38] and coherence [39, 40], the theory of quantum reference frames [41, 42], quantum clocks [43], and quantum thermodynamics [44–48]. Despite the progress that has been made, however, the full scope of the WAY theorem is still not known. For instance, the theorem as stated pertains only to sharp observables, and has been shown only in the limited framework of “normal” measurement schemes, where the apparatus is prepared in a pure state and unitarily interacts with the measured system, and where the pointer observable is sharp. But in the quantum theory of measurement [49], observables are more properly represented by positive operator valued measures (POVMs) which can be unsharp, measurement interactions are more generally described by channels which can be non-unitary, and the apparatus preparation may be described by a mixed state. Additionally, the WAY theorem addresses disturbance only in the case where the same observable is measured in succession, and the situation where the first and second observables in the sequence are not the same has received scant attention. In this paper, we shall investigate the role of conservation laws on measurement error and disturbance in the more general setting, stating our results in operational terms, in that the quantitative bounds we employ can be seen to arise from the probabilistic structure of quantum

M. Hamed Mohammady: mohammad.mohammady@ulb.be

Takayuki Miyadera: miyadera@nucleng.kyoto-u.ac.jp

Leon Loveridge: Leon.D.Loveridge@usn.no

theory in its general measurement theoretic form [50].

The paper is structured as follows. In Sec. 2, we present the elements of operational quantum theory pertinent to our investigation. This includes a background on the quantum theory of measurement, together with a quantification of measurement error and disturbance. Readers familiar with these topics can jump directly to Sec. 3, where the main results of the paper begin; here we present a framework for describing conservation laws in quantum theory, prising apart two distinct notions of conservation—full and average—whose difference manifests for general channels and which plays a key role in interpreting our findings. Next, we consider sequential measurements where the first measurement obeys a conservation law—whether average or full—and obtain general quantitative bounds for the error in the first measurement to realise a desired target observable, and the disturbance by the first measurement on a second, possibly different, observable. Here, we do not assume that the system observables are sharp, or that the apparatus pointer observable is sharp, or that the measurement interaction is unitary, or that the apparatus preparation is pure. In particular, the bounds demonstrate that in the case of a full conservation law, a large coherence in the apparatus preparation is in general a necessary condition for approximately accurate and non-disturbing measurements of observables not commuting with the system part of a conserved quantity. These bounds are then used to prove a generalisation of the WAY theorem, given in the form of a single quantitative bound, and capturing many essential features of the original theorem. Next, we provide an even stronger generalisation of the WAY theorem, indicating a deep connection between measurability, non-disturbance, and “definiteness”, and demonstrating that there are unsharp observables not commuting with the conserved quantity whose measurement cannot be accurate or non-disturbing irrespective of the apparatus preparation. Finally, in Sec. 4 we consider how the structure of the set of fixed states of the measurement channel imposes further restrictions on non-disturbance. In particular, we show that an observable not commuting with the conserved quantity admits a non-disturbing measurement only if the measurement channel disturbs all “faithful” states, i.e., states with strictly positive eigenvalues.

2 Preliminaries

In this section we introduce the elements of operational quantum theory. This includes some background on observables, instruments, and measurement schemes, as part of the quantum theory of measurement (see, e.g., [49–52]). In particular, an operationally motivated quantification of measurement error and disturbance is provided, together with a review of two special instances of non-disturbing measurements—measurements of the first kind and repeatable measurements.

2.1 Operators on Hilbert space, operations, and channels

Let \mathcal{H} be a complex separable Hilbert space, with $\mathcal{L}(\mathcal{H}) \supset \mathcal{L}_s(\mathcal{H}) \supset \mathcal{L}_p(\mathcal{H})$ the algebra of bounded (linear) operators, the real vector space of self-adjoint operators, and the (cone of) positive operators on \mathcal{H} , respectively. For any subset $\mathcal{A} \subseteq \mathcal{L}(\mathcal{H})$, the *commutant* is defined as $\mathcal{A}' := \{B \in \mathcal{L}(\mathcal{H}) : [A, B] = 0 \forall A \in \mathcal{A}\}$. We shall denote by $\mathbb{1}$ and $\mathbb{0}$ the identity and null operators of $\mathcal{L}(\mathcal{H})$, respectively, and an operator $A \in \mathcal{L}_p(\mathcal{H})$ satisfying $\mathbb{0} \leq A \leq \mathbb{1}$ will be called an *effect*. We define by $\mathcal{T}(\mathcal{H}) \subseteq \mathcal{L}(\mathcal{H})$ the two-sided ideal of trace-class operators in $\mathcal{L}(\mathcal{H})$. The (normal) state space is the space of positive, unit-trace operators $\mathcal{S}(\mathcal{H}) \subset \mathcal{T}(\mathcal{H})$, and a state $\rho \in \mathcal{S}(\mathcal{H})$ is called faithful if for all $A \in \mathcal{L}(\mathcal{H})$, $\text{tr}[A^*A\rho] = 0 \implies A = \mathbb{0}$, which implies that all of the eigenvalues of ρ are strictly positive.

Transformations of quantum systems are called *operations*, defined as completely positive (CP), trace non-increasing linear maps $\Phi : \mathcal{T}(\mathcal{H}) \rightarrow \mathcal{T}(\mathcal{K})$. Among the operations are the *channels*, which preserve the trace. For any operation $\Phi : \mathcal{T}(\mathcal{H}) \rightarrow \mathcal{T}(\mathcal{K})$, there is an associated (“Heisenberg picture”) dual operation $\Phi^* : \mathcal{L}(\mathcal{K}) \rightarrow \mathcal{L}(\mathcal{H})$, defined via the duality $\text{tr}[\Phi^*(A)T] = \text{tr}[A\Phi(T)]$ for all $A \in \mathcal{L}(\mathcal{K})$ and $T \in \mathcal{T}(\mathcal{H})$. Φ^* is completely positive and sub-unital, and unital exactly when Φ is trace-preserving. Unital operations Φ^* will also be referred to as channels. In Appendix (A) we present several properties of operations that are of central importance for the proofs of our results, most notably a Cauchy-Schwarz inequality [53].

For channels $\Phi : \mathcal{T}(\mathcal{H}) \rightarrow \mathcal{T}(\mathcal{H})$, and their duals $\Phi^* : \mathcal{L}(\mathcal{H}) \rightarrow \mathcal{L}(\mathcal{H})$, we define the fixed-point sets as

$$\mathcal{F}(\Phi) := \{T \in \mathcal{T}(\mathcal{H}) : \Phi(T) = T\}, \quad \mathcal{F}(\Phi^*) := \{A \in \mathcal{L}(\mathcal{H}) : \Phi^*(A) = A\}.$$

Note that $A \in \mathcal{F}(\Phi^*)$ is equivalent to $\text{tr}[A\Phi(\rho)] = \text{tr}[A\rho]$ for all states $\rho \in \mathcal{S}(\mathcal{H})$. Linearity of Φ^* ensures that $\mathcal{F}(\Phi^*)$ is closed under linear combinations, and because Φ^* preserves the involution, $\mathcal{F}(\Phi^*)^* = \mathcal{F}(\Phi^*)$. While $\mathcal{F}(\Phi^*)$ is not generally closed under multiplication, if $\mathcal{F}(\Phi)$ contains a faithful state then multiplicative closure is guaranteed, in which case $\mathcal{F}(\Phi^*)$ is a $*$ -algebra; in fact, it is a von Neumann algebra [54, 55]. See Appendix (B) for further details.

2.2 Observables

An *observable* of a quantum system \mathcal{S} , with Hilbert space $\mathcal{H}_\mathcal{S}$, is represented by a *normalised positive operator valued measure* (POVM) $\mathbf{E} : \Sigma \rightarrow \mathcal{L}_p(\mathcal{H}_\mathcal{S})$, where Σ is a σ -algebra of subsets of some value space \mathcal{X} , representing possible outcomes of a measurement of \mathbf{E} . For any $X \in \Sigma$, the positive operator $\mathbf{0} \leq \mathbf{E}(X) \leq \mathbf{1}_\mathcal{S}$ is referred to as an effect of \mathbf{E} . \mathbf{E} is sigma-additive on disjoint elements of Σ , and normalisation implies that $\mathbf{E}(\mathcal{X})$ is the identity operator on $\mathcal{H}_\mathcal{S}$. An effect $\mathbf{E}(X) = \alpha \mathbf{1}_\mathcal{S}$, where $\alpha \in [0, 1]$, is called *trivial*, and an observable \mathbf{E} is called non-trivial if at least one of the effects in its range is non-trivial. *Discrete observables* are those for which $\mathcal{X} = \{x_1, x_2, \dots\}$ is countable, in which case \mathbf{E} can be identified with the set $\{\mathbf{E}(x) \equiv \mathbf{E}(\{x\}) \in \mathcal{L}_p(\mathcal{H}_\mathcal{S}) : x \in \mathcal{X}\} \equiv \mathbf{E}$. If it is not stated otherwise, observables will be assumed to be discrete. Combined with states, observables give rise to the probabilities

$$p_\rho^\mathbf{E}(x) := \text{tr}[\mathbf{E}(x)\rho],$$

holding for all $\rho \in \mathcal{S}(\mathcal{H}_\mathcal{S})$ and all $x \in \mathcal{X}$, interpreted as the probability of observing outcome x when the observable \mathbf{E} is measured in the state ρ .

If \mathbf{E} is a POVM acting in $\mathcal{H}_\mathcal{S}$, the commutant of \mathbf{E} is denoted by $\mathbf{E}' := \{A \in \mathcal{L}(\mathcal{H}_\mathcal{S}) : [\mathbf{E}(x), A] = \mathbf{0} \forall x \in \mathcal{X}\}$. Since $\mathbf{E} = \mathbf{E}^*$ is a self-adjoint set, \mathbf{E}' is a von Neumann algebra, and $\mathbf{E}'' \equiv (\mathbf{E}')'$ is the smallest von Neumann algebra containing \mathbf{E} (i.e., it is the von Neumann algebra generated by \mathbf{E}). For any $A \in \mathcal{L}(\mathcal{H}_\mathcal{S})$ such that $A \in \mathbf{E}'$, we write $[\mathbf{E}, A] = \mathbf{0}$. Similarly, for any observable $\mathbf{F} := \{\mathbf{F}(y) : y \in \mathcal{Y}\}$ such that $\mathbf{F} \subset \mathbf{E}'$, we shall write $[\mathbf{E}, \mathbf{F}] = \mathbf{0}$. Among the observables are those that are *commutative*, meaning that $\mathbf{E} \subset \mathbf{E}'$ (that is, all the effects $\mathbf{E}(x)$ mutually commute). Among the commutative observables are the *sharp* observables, which satisfy the additional condition that for all $x, y \in \mathcal{X}$, $\mathbf{E}(x)\mathbf{E}(y) = \delta_{x,y}\mathbf{E}(x)$, i.e., $\mathbf{E}(x)$ are mutually orthogonal projection operators. These observables correspond to self-adjoint operators through the spectral theorem. Observables that are not sharp will be called *unsharp*, and similarly any effect E which is not a projection will be called unsharp. The unsharpness of E can be quantified through the operator norm as $0 \leq \|E - E^2\| \leq 1/4$, which vanishes exactly when E is a projection. Finally, an observable \mathbf{E} is defined as being “norm-1”, or having the norm-1 property, if $\|\mathbf{E}(x)\| = 1$ for every x for which $\mathbf{E}(x) \neq \mathbf{0}$. While sharp observables are trivially norm-1, this property may also be enjoyed by some unsharp observables.

2.3 Instruments

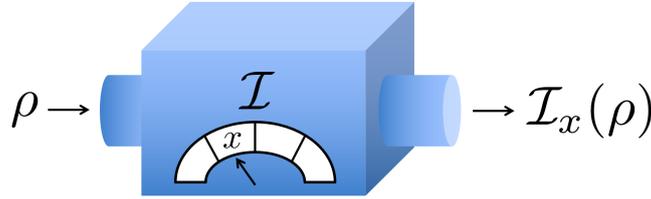


Figure 1: An instrument measures an observable \mathbf{E} of the system \mathcal{S} , and also transforms the system conditional on registering a given outcome. The system, initially prepared in an arbitrary state ρ , enters the instrument which then registers outcome x with probability $p_\rho^\mathbf{E}(x) := \text{tr}[\mathbf{E}(x)\rho] = \text{tr}[\mathcal{I}_x(\rho)]$. Subsequently, the instrument transforms the system to the (non-normalised) state $\mathcal{I}_x(\rho)$.

Though the state-observable pairings describe the totality of the measurement statistics, this is not sufficient for determining other interesting properties of a measurement, for instance the form of the associated state changes. To this end, we make use of the notion of *instrument*, or *operation valued measure* [56–60]. A discrete instrument is a collection of operations $\mathcal{I} := \{\mathcal{I}_x \equiv \mathcal{I}_{\{x\}} : x \in \mathcal{X}\}$ such that $\mathcal{I}_\mathcal{X}(\cdot) := \sum_{x \in \mathcal{X}} \mathcal{I}_x(\cdot)$ is a channel. Throughout, we shall always assume that \mathcal{I} acts in $\mathcal{H}_\mathcal{S}$, that is, $\mathcal{I}_x : \mathcal{T}(\mathcal{H}_\mathcal{S}) \rightarrow \mathcal{T}(\mathcal{H}_\mathcal{S})$. Each instrument is associated with a unique observable \mathbf{E} via $\mathcal{I}_x^*(\mathbf{1}_\mathcal{S}) = \mathbf{E}(x)$, which implies that $p_\rho^\mathbf{E}(x) := \text{tr}[\mathbf{E}(x)\rho] = \text{tr}[\mathcal{I}_x(\rho)]$. We refer to such an \mathcal{I} as an \mathbf{E} -compatible instrument, or an \mathbf{E} -instrument for short, and to $\mathcal{I}_\mathcal{X}(\cdot)$ as the associated \mathbf{E} -channel. $\mathcal{I}_x(\rho)$ is interpreted as the non-normalised state after a measurement of \mathbf{E} has taken place and the outcome x has been registered, and $\mathcal{I}_\mathcal{X}(\rho)$ is the normalised state after a non-selective measurement. A schematic representation of an instrument is given in Fig. 1.

We note that for every discrete observable \mathbf{E} , there are infinitely many \mathbf{E} -compatible instruments; every \mathbf{E} -instrument \mathcal{I} can be constructed as the set of operations $\{\mathcal{I}_x = \Phi_x \circ \mathcal{I}_x^L : x \in \mathcal{X}\}$ [58, 60], where $\Phi_x : \mathcal{T}(\mathcal{H}_\mathcal{S}) \rightarrow \mathcal{T}(\mathcal{H}_\mathcal{S})$ are arbitrary channels that may depend on outcome x , and \mathcal{I}^L is the Lüders instrument for \mathbf{E} [61], defined as

$$\mathcal{I}_x^L(T) := \sqrt{\mathbf{E}(x)}T\sqrt{\mathbf{E}(x)} \quad \mathcal{I}_x^{L*}(A) := \sqrt{\mathbf{E}(x)}A\sqrt{\mathbf{E}(x)}, \quad (1)$$

to hold for all $x \in \mathcal{X}$, $T \in \mathcal{T}(\mathcal{H}_\mathcal{S})$, and $A \in \mathcal{L}(\mathcal{H}_\mathcal{S})$.

2.4 Measurement schemes

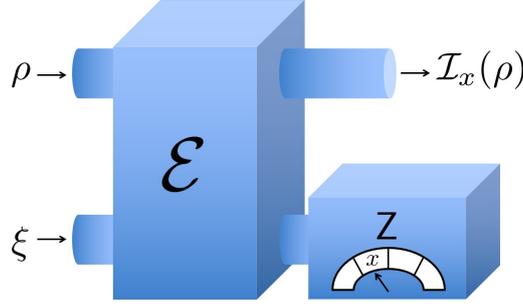


Figure 2: An E-instrument \mathcal{I} is implemented on the system \mathcal{S} via a measurement scheme. The system, initially prepared in an arbitrary state ρ , and a measuring apparatus \mathcal{A} , prepared in a fixed state ξ , undergo a joint evolution by the channel \mathcal{E} . Subsequently, a pointer observable Z of the apparatus is measured. With probability $p_\rho^E(x) := \text{tr}[\mathbf{E}(x)\rho] = \text{tr}[\mathcal{I}_x(\rho)]$ the apparatus registers outcome x , thereby transforming the system to the non-normalised state $\mathcal{I}_x(\rho)$.

An even more comprehensive description of the measurement process involves the modelling of a measuring apparatus \mathcal{A} and a specification of how it couples to the system under investigation. A *measurement scheme* is characterised by the tuple $\mathcal{M} := (\mathcal{H}_A, \xi, \mathcal{E}, Z)$ where: \mathcal{H}_A is the Hilbert space for the measuring apparatus \mathcal{A} and $\xi \in \mathcal{S}(\mathcal{H}_A)$ is a state on \mathcal{H}_A ; $\mathcal{E} : \mathcal{T}(\mathcal{H}_S \otimes \mathcal{H}_A) \rightarrow \mathcal{T}(\mathcal{H}_S \otimes \mathcal{H}_A)$ is a channel which serves to correlate system and apparatus; and $Z := \{Z(x) : x \in \mathcal{X}\}$ is a “pointer” observable of the apparatus. The operations of the instrument implemented by \mathcal{M} can be written as

$$\mathcal{I}_x(\cdot) = \text{tr}_A[(\mathbf{1}_S \otimes Z(x))\mathcal{E}(\cdot \otimes \xi)], \quad (2)$$

where $\text{tr}_A : \mathcal{T}(\mathcal{H}_S \otimes \mathcal{H}_A) \rightarrow \mathcal{T}(\mathcal{H}_S)$ is the partial trace channel over the apparatus, defined as $\text{tr}[A \text{tr}_A[T]] = \text{tr}[(A \otimes \mathbf{1}_A)T]$ for all $A \in \mathcal{L}(\mathcal{H}_S)$ and $T \in \mathcal{T}(\mathcal{H}_S \otimes \mathcal{H}_A)$. The channel implemented by \mathcal{M} may thus be written as $\mathcal{I}_\mathcal{X}(\cdot) = \text{tr}_A[\mathcal{E}(\cdot \otimes \xi)]$. A schematic representation of a measurement scheme is given in Fig. 2. We note that every E-compatible instrument admits infinitely many *normal* measurement schemes, where ξ is chosen to be pure, \mathcal{E} is chosen to be unitary, and Z is chosen to be sharp [62]. However, unless stated otherwise, we shall consider the more general situation where ξ may be mixed, \mathcal{E} may be non-unitary, and Z may be unsharp.

We now introduce the unital, completely positive normal conditional expectation $\Gamma_\xi : \mathcal{L}(\mathcal{H}_S \otimes \mathcal{H}_A) \rightarrow \mathcal{L}(\mathcal{H}_S)$. Γ_ξ , called a *restriction map* for ξ , is defined as the dual of the isometric embedding (or the *preparation map*) $T \mapsto T \otimes \xi$, and satisfies $\text{tr}[\Gamma_\xi(B)T] = \text{tr}[B(T \otimes \xi)]$ for all $B \in \mathcal{L}(\mathcal{H}_S \otimes \mathcal{H}_A)$ and $T \in \mathcal{T}(\mathcal{H}_S)$. We may use the restriction map to define the channel $\Gamma_\xi^\mathcal{E} : \mathcal{L}(\mathcal{H}_S \otimes \mathcal{H}_A) \rightarrow \mathcal{L}(\mathcal{H}_S)$ as

$$\Gamma_\xi^\mathcal{E}(\cdot) := \Gamma_\xi \circ \mathcal{E}^*(\cdot). \quad (3)$$

Using Eq. (3), we may express the duals of the operations defined in Eq. (2) as

$$\mathcal{I}_x^*(\cdot) = \Gamma_\xi^\mathcal{E}(\cdot \otimes Z(x)). \quad (4)$$

In particular, we may write the dual channel as $\mathcal{I}_\mathcal{X}^*(\cdot) = \Gamma_\xi^\mathcal{E}(\cdot \otimes \mathbf{1}_A)$.

We may also be interested in asking how the apparatus is transformed as a result of the measurement interaction. To this end, we introduce the channel $\Lambda : \mathcal{T}(\mathcal{H}_S) \rightarrow \mathcal{T}(\mathcal{H}_A)$ and its dual $\Lambda^* : \mathcal{L}(\mathcal{H}_A) \rightarrow \mathcal{L}(\mathcal{H}_S)$, referred to as *conjugate* channels to $\mathcal{I}_\mathcal{X}$ and $\mathcal{I}_\mathcal{X}^*$, respectively, defined as

$$\Lambda(T) := \text{tr}_S[\mathcal{E}(T \otimes \xi)], \quad \Lambda^*(A) := \Gamma_\xi^\mathcal{E}(\mathbf{1}_S \otimes A), \quad (5)$$

to hold for all $T \in \mathcal{L}(\mathcal{H}_S)$ and $A \in \mathcal{L}(\mathcal{H}_A)$. That is, $\Lambda(\rho)$ is the state of the apparatus after it has interacted with the system, when the system is initially prepared in state ρ . On the other hand, for an initial system state ρ , the expected value of $A \in \mathcal{L}(\mathcal{H}_A)$ in the state of the apparatus after the measurement interaction can be obtained by evaluating the expected value of $\Lambda^*(A)$ in ρ .

2.5 Quantifying measurement error and measurement disturbance

In order to quantify *measurement error* and *measurement disturbance*, we shall first provide a quantification of the difference, or discrepancy, between two effects E and F acting in a generic space \mathcal{H} . For any state $\rho \in \mathcal{S}(\mathcal{H})$, the probabilities that the properties corresponding to E and F are realised can be compared as $|\text{tr}[\rho(E - F)]|$,

which can be estimated through repeated measurements of E and F in the state ρ . Given that we wish to quantify the sense in which E and F differ as effects, i.e., independently of the state, it is natural to take the supremum over all states, and note that

$$\sup_{\rho \in \mathcal{S}(\mathcal{H})} |\text{tr}[\rho(E - F)]| = \|E - F\|, \quad (6)$$

the right hand side denoting the operator norm, which of course vanishes when $E = F$ and is non-zero otherwise. Eq. (6) gives an operationally motivated—in the sense of being derived directly from the experimental probabilities—quantification of the discrepancy between two effects, which will be utilised in the analysis of measurement error and disturbance.

Let us first address the question of measurement error. Note that by Eq. (4) and Eq. (5), the observable that is measured by a scheme $\mathcal{M} := (\mathcal{H}_A, \xi, \mathcal{E}, \mathbf{Z})$ has the effects $\mathcal{I}_x^*(\mathbb{1}_S) = \Gamma_\xi^\mathcal{E}(\mathbb{1}_S \otimes \mathbf{Z}(x)) \equiv \Lambda^*(\mathbf{Z}(x))$. Now, let \mathbf{E} be the *target* observable, i.e., the observable we wish to measure, which may be different to the measured observable, but has the same value space \mathcal{X} . By Eq. (6), the measurement error for each effect of the target observable can be quantified through the operator norm as $\|\epsilon(x)\|$, where

$$\epsilon(x) := \Lambda^*(\mathbf{Z}(x)) - \mathbf{E}(x). \quad (7)$$

A global quantification of measurement error may thus be defined as the largest error over all effects of \mathbf{E} ,

$$\epsilon := \max_{x \in \mathcal{X}} \|\epsilon(x)\|.$$

\mathcal{M} is a measurement scheme for \mathbf{E} if $\epsilon = 0$, that is, if the target observable is what is actually measured, so that $\Lambda^*(\mathbf{Z}(x)) = \mathbf{E}(x)$ for all x . In the absence of any constraints, perfectly accurate measurements are possible for any target observable \mathbf{E} .

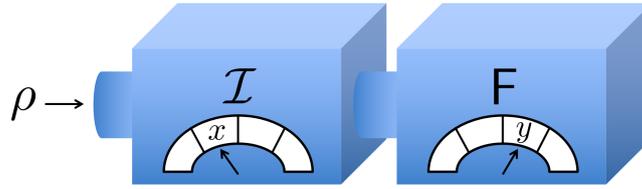


Figure 3: In a sequential measurement, an observable \mathbf{E} is measured by an instrument \mathcal{I} , and subsequently a second observable \mathbf{F} is measured. \mathcal{I} does not disturb \mathbf{F} if for all input states ρ , the statistics of \mathbf{F} do not depend on whether an \mathbf{E} -measurement took place or not.

Now let us address the question of measurement disturbance. Let $\mathbf{E} := \{\mathbf{E}(x) : x \in \mathcal{X}\}$ and $\mathbf{F} := \{\mathbf{F}(y) : y \in \mathcal{Y}\}$ be two observables acting in \mathcal{H}_S . Consider the sequential measurement of these observables, as depicted in Fig. 3, where at first \mathbf{E} is measured by the instrument \mathcal{I} (implemented by some measurement scheme \mathcal{M}), and subsequently \mathbf{F} is measured. For any initial state $\rho \in \mathcal{S}(\mathcal{H}_S)$, the probability of observing outcome y of \mathbf{F} after a non-selective measurement by the \mathbf{E} -instrument \mathcal{I} is given as

$$\text{tr}[\mathbf{F}(y)\mathcal{I}_X(\rho)] \equiv \text{tr}[\mathcal{I}_X^*(\mathbf{F}(y))\rho].$$

That is, the prior \mathbf{E} -measurement implies that we perform a measurement of the disturbed observable $\{\mathcal{I}_X^*(\mathbf{F}(y)) : y \in \mathcal{Y}\}$ in the state ρ . By Eq. (6), the disturbance of each effect of \mathbf{F} may be quantified through the operator norm as $\|\delta(y)\|$, where

$$\delta(y) := \mathcal{I}_X^*(\mathbf{F}(y)) - \mathbf{F}(y). \quad (8)$$

Note that if \mathcal{I} is implemented by the measurement scheme $\mathcal{M} := (\mathcal{H}_A, \xi, \mathcal{E}, \mathbf{Z})$, then we may equivalently write $\delta(y) = \Gamma_\xi^\mathcal{E}(\mathbf{F}(y) \otimes \mathbb{1}_A) - \mathbf{F}(y)$. A global quantification of the disturbance of \mathbf{F} can then be defined as the largest disturbance over all the effects,

$$\delta := \max_{y \in \mathcal{Y}} \|\delta(y)\|,$$

and \mathbf{F} is non-disturbed by \mathcal{I} exactly when $\delta = 0$, which is the case when $\mathcal{I}_X^*(\mathbf{F}(y)) = \mathbf{F}(y)$ for all $y \in \mathcal{Y}$. In other words, \mathcal{I} does not disturb \mathbf{F} exactly when each $\mathbf{F}(y)$ is a fixed point of the \mathbf{E} -channel \mathcal{I}_X^* , i.e., $\mathbf{F} \subset \mathcal{F}(\mathcal{I}_X^*)$. In such a case, for any initial state ρ , a non-selective measurement by \mathcal{I} does not affect the subsequent measurement statistics of \mathbf{F} . In the absence of any constraints, non-disturbance is always possible when the pair of observables

commute; since $E' \subset \mathcal{F}(\mathcal{I}_\lambda^{L*})$ always holds, where \mathcal{I}^L is the Lüders E-instrument defined in Eq. (1), then a Lüders measurement of E is guaranteed not to disturb any F commuting with E [63]. While the fixed-point set of the E-channel \mathcal{I}_λ^* is not always contained in the commutant of E, in Appendix (C) we present some cases where $\mathcal{F}(\mathcal{I}_\lambda^*) \subset E'$ necessarily holds. For a wider discussion on the relationship between disturbance, commutation, and compatibility, and a quantitative bound relating the minimum disturbance in terms of the commutation between the pair of observables, and the unsharpness of each, see Appendix (D).

2.6 Measurements of the first kind, and repeatable measurements

A special instance of a non-disturbing measurement is when an E-instrument \mathcal{I} does not disturb E itself, i.e., when $E \subset \mathcal{F}(\mathcal{I}_\lambda^*)$. Such measurements are referred to as *measurements of the first kind*. A subclass of measurements of the first kind are those which are *repeatable*. Though repeatability is a standard assumption in many textbook treatments of quantum mechanics, that it is a property which a measurement may or may not enjoy appeared already in Wigner’s 1952 contribution on the WAY theorem. However, within the general framework presented thus far, repeatability corresponds to a very special form of state change, possible only for a privileged class of observables—an observable E admits a repeatable measurement only if it is discrete [62], and all the effects have at least one eigenvector with eigenvalue 1 [64]. \mathcal{I} is a repeatable E-instrument if

$$\mathrm{tr}[\mathcal{I}_y \circ \mathcal{I}_x(\rho)] = \delta_{x,y} \mathrm{tr}[\mathcal{I}_x(\rho)] \quad \forall \rho \in \mathcal{S}(\mathcal{H}_S), x, y \in \mathcal{X},$$

which implies that

$$\mathcal{I}_x^*(E(y)) = \delta_{x,y} E(x) \quad \forall x, y \in \mathcal{X}. \quad (9)$$

The above definition is equivalent to $\mathcal{I}_x^*(E(x)) = E(x)$ for all x , since if this holds then $\mathcal{I}_x^*(\mathbb{1}_S - E(x)) = E(x) - E(x) = \mathbb{O}$ [1]. In other words, if \mathcal{I} is a repeatable instrument, then repeated measurements by \mathcal{I} are guaranteed (with probability one) to produce the same result. It is straightforward to verify that if a measurement of E is repeatable, then it is also of the first kind, since

$$\mathcal{I}_\lambda^*(E(y)) = \sum_{x \in \mathcal{X}} \mathcal{I}_x^*(E(y)) = \sum_{x \in \mathcal{X}} \delta_{x,y} E(x) = E(y).$$

While the converse relation does not hold in general—a measurement can be of the first kind and not repeatable, such as is the case for a Lüders instrument compatible with a commutative but unsharp observable—in the special case of sharp observables repeatability and first-kindness coincide (Theorem 1 in Ref. [65]). In Appendix (E), we provide a series of results regarding the structure of repeatable instruments.

3 Generalisation of the Wigner-Araki-Yanase theorem

The Wigner-Araki-Yanase (WAY) theorem is the classic result connecting measurement, conservation, and disturbance. This theorem was formulated by Araki and Yanase in 1960 [25], capturing in a fairly general setting an observation due to Wigner given in 1952 [23, 24] regarding spin measurements in the presence of angular momentum conservation. The WAY theorem as formulated in Ref. [26] states that for any discrete sharp observable represented as a self-adjoint operator A not commuting with the system part of a (bounded, additive) conserved quantity, the measurement—described by a normal measurement scheme—cannot be repeatable and must violate the Yanase condition, i.e., the pointer observable of the apparatus must fail to commute with the apparatus part of the conserved quantity. In other words, if the Yanase condition is satisfied, then the measurement cannot be “accurate”, in the sense that A is not measured by the scheme. But the WAY theorem does not rule out approximate measurements with approximate repeatability properties, where approximate measurement is understood to mean that the unsharp observable which is actually measured can be made statistically close to A . Therefore, WAY has both a strict impossibility part, along with the provision of conditions under which approximate measurements may be possible; as already hinted at by Wigner’s original observation, and subsequently refined by Yanase [66] and Ozawa [67] in the form of quantitative bounds, a normal measurement scheme obeying a conservation law and the Yanase condition can achieve approximately accurate measurements for A only if the uncertainty of the apparatus preparation in the conserved quantity is large.

While the WAY theorem has developed over the years, its full scope is still not known. In particular, while much of the previous work around the WAY theorem has focused on the “measurability question”—upon which observables cannot be measured, or can be only approximately measured given the conservation law—the role of

disturbance has been much less fully examined. Moreover, previous proofs of the WAY theorem concerned only sharp target observables, and were shown in the limited framework of normal measurement schemes, leaving open the question as to whether the implications of the theorem will carry over to the more general setting. In this section, we shall close this gap.

We begin by first introducing two operational definitions of conservation laws for channels—full conservation, and the generally weaker notion of average conservation. Next, we obtain quantitative bounds for measurement error and measurement disturbance in the presence of conservation—whether average or full—in the general setting, i.e., without assuming sharpness of the measured observable, the sharpness of the pointer observable, the unitarity of the interaction or the purity of the apparatus preparation. These bounds allow us to prove a generalisation of WAY, which is presented as a single quantitative bound that contains both the possibility and impossibility statements of the theorem. A further generalisation of the WAY theorem is also provided, this time presented as an equality that serves to further strengthen the strict impossibility statement of the WAY theorem for observables that may be unsharp, but are “definite”. The section concludes with a demonstration that by imposing conservation laws on pointer objectification in addition to the measurement interaction between system and apparatus, the measurability part of WAY can be recovered without the Yanase condition.

3.1 Measurement schemes in the presence of conservation laws

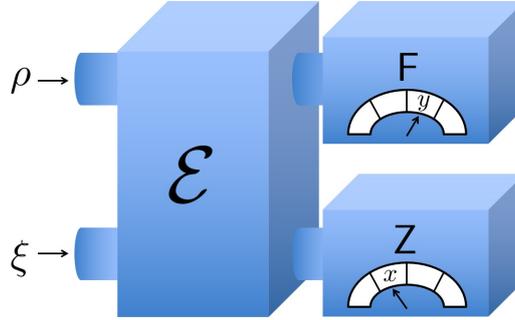


Figure 4: The system is measured in succession, with the first measurement implemented by a scheme $\mathcal{M} := (\mathcal{H}_A, \xi, \mathcal{E}, Z)$ where \mathcal{E} conserves an additive quantity $N = N_S \otimes \mathbb{1}_A + \mathbb{1}_S \otimes N_A$. The conservation law imposes restrictions on the possibility of \mathcal{M} to accurately measure a target observable E , and on the possibility of non-disturbance for the second observable F .

While every E -compatible instrument \mathcal{I} admits some measurement scheme $\mathcal{M} := (\mathcal{H}_A, \xi, \mathcal{E}, Z)$, any restrictions imposed on the elements of \mathcal{M} will in turn restrict the types of instruments that can be implemented, and hence the class of observables E that can be accurately measured, and the class of observables F that will be non-disturbed. One such restriction is given by conservation laws—for example, the interaction channel \mathcal{E} between system and apparatus may be restricted so that the total energy, charge, or angular momentum must be conserved. Before investigating how conservation limits measurements, let us first consider two operational definitions of conservation laws for channels, where the conserved quantity N is always assumed to be a bounded self-adjoint operator. In the first analysis, a conservation law can be defined by equality of expectation values before and after the action of the channel, i.e., average conservation:

Definition 1. A channel $\Phi : \mathcal{T}(\mathcal{H}) \rightarrow \mathcal{T}(\mathcal{H})$ conserves a self-adjoint operator $N \in \mathcal{L}_s(\mathcal{H})$ on average if for all $\rho \in \mathcal{S}(\mathcal{H})$,

$$\text{tr}[N\Phi(\rho)] = \text{tr}[N\rho],$$

i.e., $N \in \mathcal{F}(\Phi^*)$.

However, this does not rule out the higher moments of the “conserved” quantity changing their values. Thus we may strengthen the definition in the following way:

Definition 2. A channel $\Phi : \mathcal{T}(\mathcal{H}) \rightarrow \mathcal{T}(\mathcal{H})$ fully conserves a self-adjoint operator $N \in \mathcal{L}_s(\mathcal{H})$ if for all $\rho \in \mathcal{S}(\mathcal{H})$ and $k \in \mathbb{N}$,

$$\text{tr}[N^k\Phi(\rho)] = \text{tr}[N^k\rho],$$

i.e., $N^k \in \mathcal{F}(\Phi^*)$ for all $k \in \mathbb{N}$.

As shown in Appendix (F), full conservation is in fact equivalent to just the first two moments being conserved, i.e., $N^k \in \mathcal{F}(\Phi^*)$ for $k = 1, 2$. Moreover, full conservation is also shown to be equivalent to “invariance” of the unitary group generated by N under the action of the channel, i.e., $\Phi^*(e^{itN}) = e^{itN}$ for all $t \in \mathbb{R}$. We

note that invariance implies (but is not equivalent to) ‘‘covariance’’, i.e., $\Phi^*(e^{itN} A e^{-itN}) = e^{itN} \Phi^*(A) e^{-itN}$ for all $t \in \mathbb{R}$ and $A \in \mathcal{L}(\mathcal{H})$. While full conservation trivially implies average conservation, however, it is shown that in general a channel may conserve N on average but not fully. Indeed, it is possible for a channel to conserve N on average while not being covariant. Therefore, average conservation is generally a weaker form of conservation law, and is logically distinct from the concept of ‘‘symmetry’’ [68, 69]. However, in the special case where $\Phi(\cdot) := U \cdot U^*$ is a unitary channel, average and full conservation coincide, and are both equivalent to the commutation relation $[U, N] = \mathbf{O}$. Since a normal measurement scheme uses a unitary interaction channel, it follows that in such cases there is no distinction to be drawn between the two notions of conservation law. But if a measurement scheme is not normal, i.e., if the interaction channel is non-unitary, then the distinction between average and full conservation will no longer be void and, as we shall see, leads to interesting consequences.

Throughout what follows, we shall only consider the case where the interaction channel \mathcal{E} conserves a quantity N that is a bounded, additive, self-adjoint operator. That is, $N = N_S \otimes \mathbf{1}_A + \mathbf{1}_S \otimes N_A$, where $N_S \in \mathcal{L}_s(\mathcal{H}_S)$ and $N_A \in \mathcal{L}_s(\mathcal{H}_A)$ are respectively bounded quantities of the system and apparatus alone. Note that conservation of an additive N by the interaction channel \mathcal{E} does not generally imply conservation of N_S by the channel \mathcal{I}_X , since \mathcal{E} may allow for an ‘‘exchange’’ of the conserved quantity between system and apparatus; specifically, average conservation of N by \mathcal{E} implies that

$$\text{tr}[N_S(\mathcal{I}_X(\rho) - \rho)] = \text{tr}[N_A(\xi - \Lambda(\rho))] \quad (10)$$

holds for all $\rho \in \mathcal{S}(\mathcal{H}_S)$, where Λ is the conjugate channel of \mathcal{I}_X defined in Eq. (5), so that $\Lambda(\rho)$ is the state of the apparatus after the measurement interaction. We can see that it is possible for the expected value of N_S to increase (decrease), provided that the expected value of N_A decreases (increases) by an equal amount. Indeed, such a ‘‘compensation’’ by the measuring apparatus is in general necessary for the instrument \mathcal{I} to accurately measure some observable E not commuting with the conserved quantity: if E is a sharp observable and the E -channel \mathcal{I}_X conserves N_S on average, by item (i) of Lemma C.1 it holds that E must commute with N_S . Additionally, if \mathcal{I}_X fully conserves N_S then by item (iii) of Lemma C.1 E must commute with N_S , even when E is unsharp.

3.2 Measurement error and disturbance under conservation laws

Consider the case where the system is measured in succession, but where the first measurement is constrained by a conservation law, as shown in Fig. 4. We now present our first main result, providing quantitative bounds for the error of the first measurement in achieving the desired target observable E , and for the disturbance caused by the first measurement on the second observable F . These bounds will be used to obtain our generalisation of the WAY theorem in the sequel.

Theorem 3.1. *Let $\mathcal{M} := (\mathcal{H}_A, \xi, \mathcal{E}, \mathcal{Z})$ be a measurement scheme for an instrument $\mathcal{I} := \{\mathcal{I}_x : x \in \mathcal{X}\}$ acting in \mathcal{H}_S , and assume that \mathcal{E} conserves an additive quantity $N = N_S \otimes \mathbf{1}_A + \mathbf{1}_S \otimes N_A$ on average, where $N_S \in \mathcal{L}_s(\mathcal{H}_S)$ and $N_A \in \mathcal{L}_s(\mathcal{H}_A)$. Let $\|\epsilon(x)\|$ be the error in measuring the effects of a target observable $E := \{E(x) : x \in \mathcal{X}\}$ by \mathcal{M} , as defined in Eq. (7). Then for all $x \in \mathcal{X}$ it holds that*

$$\| [E(x), N_S] - \Lambda^*([Z(x), N_A]) \| \leq 2\|N_S\| \|\epsilon(x)\| + 2\|\Gamma_\xi^\mathcal{E}(N^2) - \Gamma_\xi^\mathcal{E}(N)^2\|^{\frac{1}{2}} \left(2\|\epsilon(x)\| + \|E(x) - E(x)^2\| \right)^{\frac{1}{2}}, \quad (11)$$

where $\Gamma_\xi^\mathcal{E}$ is the channel defined in Eq. (3), and Λ is the conjugate channel to \mathcal{I}_X defined in Eq. (5).

Let $\|\delta(y)\|$ be the disturbance of the effects of an observable $F := \{F(y) : y \in \mathcal{Y}\}$ caused by \mathcal{I} , as defined in Eq. (8). Then for all $y \in \mathcal{Y}$ it holds that

$$\| [F(y), N_S] - \mathcal{I}_X^*([F(y), N_S]) \| \leq 2\|N_S\| \|\delta(y)\| + 2\|\Gamma_\xi^\mathcal{E}(N^2) - \Gamma_\xi^\mathcal{E}(N)^2\|^{\frac{1}{2}} \left(2\|\delta(y)\| + \|F(y) - F(y)^2\| \right)^{\frac{1}{2}}. \quad (12)$$

The proof for the error bound Eq. (11) is provided in Appendix (G), and the proof for the disturbance bound Eq. (12) is given in Appendix (H). In Appendix (D), we also provide similar bounds for disturbance that are independent of conservation, but take into account the commutation between F and the observable that is measured by \mathcal{I} . Note that while the upper bounds of both inequalities above (the terms on the right hand side) are structurally the same, the lower bounds (the terms on the left hand side) are not. Specifically, while the constraints on measurement error depend on the choice of pointer observable, the constraints on disturbance are independent of this. It follows that the implications of these inequalities differ markedly. We shall illustrate this by considering when the inequalities impose no constraints, i.e., when the lower bounds vanish.

Let us first consider Eq. (11). This will not impose any constraints on perfectly accurate measurements, i.e., $\epsilon = 0$, if it holds that $\Lambda^*([Z(x), N_A]) = [E(x), N_S]$, as in such a case the lower bound vanishes. If the pointer observable obeys the Yanase condition, i.e., if Z commutes with N_A , it follows that there will not be any constraints on perfectly accurate measurements if E commutes with N_S , since by complete positivity it holds that $[Z(x), N_A] = [E(x), N_S] = \mathbf{0} \implies \Lambda^*([Z(x), N_A]) = [E(x), N_S]$. But if we are free to choose the pointer observable, then there exists a measurement scheme obeying a conservation law that can accurately measure all target observables E . For example, let us assume that \mathcal{M} is a “trivial” measurement scheme where the system and apparatus are identical, and where \mathcal{E} is a unitary swap channel. That is, choose $\mathcal{H}_A \simeq \mathcal{H}_S$, $N_A = N_S$, and $\mathcal{E}^*(A \otimes B) = B \otimes A$ for all A, B . It is clear that \mathcal{E} conserves $N_S \otimes \mathbf{1}_A + \mathbf{1}_S \otimes N_S$. In such a case, $\Lambda^*(A) = A$ for all A , and so by choosing the pointer observable so that $Z = E$, we obtain $\epsilon(x) = \Lambda^*(E(x)) - E(x) = \mathbf{0}$, and so all target observables are measurable. This is perfectly consistent with Eq. (11), since in such a case $\Lambda^*([E(x), N_S]) = [E(x), N_S]$, and so the lower bound vanishes. But note that if \mathcal{M} is trivial, then the instrument that it implements is also trivial, i.e., it will hold that $\mathcal{I}_{\mathcal{X}}^*(\cdot) = \text{tr}[\cdot\xi]\mathbf{1}_S$. In such a case, we have $\delta(y) = \text{tr}[F(y)\xi]\mathbf{1}_S - F(y)$, so all non-trivial observables will be disturbed.

Let us now consider Eq. (12). This will not impose any constraints on non-disturbance for F , i.e., $\delta = 0$, if F commutes with the conserved quantity. This is because by complete positivity, it holds that $[F(y), N_S] = \mathbf{0} \implies \mathcal{I}_{\mathcal{X}}^*([F(y), N_S]) = [F(y), N_S]$, in which case the lower bound vanishes. But note that $\mathcal{I}_{\mathcal{X}}^*([F(y), N_S]) = [F(y), N_S]$ does not in general imply that $[F(y), N_S] = \mathbf{0}$. For example, let us assume that $\mathcal{I}_{\mathcal{X}}$ fully conserves N_S , i.e., $N_S^k \in \mathcal{F}(\mathcal{I}_{\mathcal{X}}^*)$ for $k = 1, 2$. In such a case, by the multiplicability theorem (Corollary A.1) it will hold that $\mathcal{I}_{\mathcal{X}}^*([A, N_S]) = [\mathcal{I}_{\mathcal{X}}^*(A), N_S]$ for all $A \in \mathcal{L}(\mathcal{H}_S)$. If it additionally holds that $F \in \mathcal{F}(\mathcal{I}_{\mathcal{X}}^*)$, then $\mathcal{I}_{\mathcal{X}}^*([F(y), N_S]) = [F(y), N_S]$, so that the lower bound of Eq. (12) will vanish and non-disturbance will not be ruled out for F , whether it commutes with N_S or not. But by item (iii) of Lemma C.1 such an instrument \mathcal{I} will accurately measure E only if $[E, N_S] = \mathbf{0}$. Indeed, as a result of the above arguments, and as shown in Corollary H.2, if the measurement scheme \mathcal{M} implements the Lüders instrument \mathcal{I}^L compatible with an observable E commuting with N_S , then non-disturbance will not be ruled out for any observable F that commutes with E .

Notwithstanding the special cases discussed above, when E and F do not commute with the conserved quantity, the lower bounds in Theorem 3.1 will not vanish in general, in which case the upper bounds must also not vanish. It follows that a large value of $\|\Gamma_{\xi}^{\mathcal{E}}(N^2) - \Gamma_{\xi}^{\mathcal{E}}(N)^2\|$ is a necessary condition for achieving an arbitrarily small measurement error for E when $[E(x), N_S] \neq \Lambda^*([Z(x), N_A])$ and an arbitrarily small disturbance for F when $[F(y), N_S] \neq \mathcal{I}_{\mathcal{X}}^*([F(y), N_S])$. If the error and disturbance are to be exactly zero, then E and F must also be unsharp. The term $\|\Gamma_{\xi}^{\mathcal{E}}(N^2) - \Gamma_{\xi}^{\mathcal{E}}(N)^2\|$ is clearly dependent on the choice of apparatus preparation ξ and, as we show below, under the stronger constraint of a full conservation law this quantity obtains a clearer interpretation as the uncertainty of N_A in the apparatus preparation, as quantified by the variance.

Lemma 3.1. *If the channel \mathcal{E} fully conserves an additive quantity $N = N_S \otimes \mathbf{1}_A + \mathbf{1}_S \otimes N_A$, then*

$$\|\Gamma_{\xi}^{\mathcal{E}}(N^2) - \Gamma_{\xi}^{\mathcal{E}}(N)^2\| = \text{Var}(N_A, \xi),$$

where $\text{Var}(N_A, \xi) := \text{tr}[N_A^2\xi] - \text{tr}[N_A\xi]^2$ denotes the variance of N_A in the state ξ .

Proof. If N is fully conserved by \mathcal{E} , then by Definition 2 we have $\mathcal{E}^*(N^k) = N^k$ for $k = 1, 2$. It follows that $\Gamma_{\xi}^{\mathcal{E}}(N^k) = \Gamma_{\xi}(N^k)$ for $k = 1, 2$, and hence $\Gamma_{\xi}^{\mathcal{E}}(N^2) - \Gamma_{\xi}^{\mathcal{E}}(N)^2 = \Gamma_{\xi}(N^2) - \Gamma_{\xi}(N)^2$. Recall that the restriction map satisfies $\Gamma_{\xi}(A \otimes B) = \text{tr}[B\xi]A$ for all $A \in \mathcal{L}(\mathcal{H}_S)$ and $B \in \mathcal{L}(\mathcal{H}_A)$. It follows that $\Gamma_{\xi}(N) = N_S + \text{tr}[N_A\xi]\mathbf{1}_S$. Now note that $N^2 = N_S^2 \otimes \mathbf{1}_A + 2N_S \otimes N_A + \mathbf{1}_S \otimes N_A^2$. Therefore,

$$\begin{aligned}\Gamma_{\xi}(N^2) &= N_S^2 + 2\text{tr}[N_A\xi]N_S + \text{tr}[N_A^2\xi]\mathbf{1}_S, \\ \Gamma_{\xi}(N)^2 &= N_S^2 + 2\text{tr}[N_A\xi]N_S + \text{tr}[N_A\xi]^2\mathbf{1}_S,\end{aligned}$$

which gives

$$\Gamma_{\xi}(N^2) - \Gamma_{\xi}(N)^2 = \text{tr}[N_A^2\xi]\mathbf{1}_S - \text{tr}[N_A\xi]^2\mathbf{1}_S = \text{Var}(N_A, \xi)\mathbf{1}_S.$$

□

If ξ is a pure state, then a large variance implies a large coherence. This is because $\text{Var}(N_A, \xi) = 0 \iff [N_A, \xi] = \mathbf{0}$ for pure states. Of course, if ξ is a mixed state then it may still be the case that $\text{Var}(N_A, \xi)$ is large even if ξ commutes with N_A , and hence has zero coherence in the conserved quantity. A quantifier of coherence (or asymmetry) for general states is given by the quantum Fisher information [70–73], which is equal to four times the convex roof of the variance [74, 75]. Let $\{q_i, \phi_i\}$ be an arbitrary ensemble of (not necessarily

orthogonal) unit vectors $\phi_i \in \mathcal{H}_A$, with $\{q_i\}$ a probability distribution. The quantum Fisher information of N_A in ξ can be written as

$$\mathcal{Q}(N_A, \xi) = 4 \inf_{\{q_i, \phi_i\}} \left\{ \sum_i q_i \text{Var}(N_A, \phi_i) : \xi = \sum_i q_i P_{\phi_i} \right\}. \quad (13)$$

Here, $P_\psi \equiv |\psi\rangle\langle\psi|$ denotes the projection on ψ , and we use the short-hand notation $\text{Var}(N_A, \phi_i) \equiv \text{Var}(N_A, P_{\phi_i})$. It is clear that $\mathcal{Q}(N_A, \xi) = 0$ if and only if $[N_A, \xi] = \mathbf{0}$, while $\mathcal{Q}(N_A, \xi) = 4\text{Var}(N_A, \xi)$ if ξ is a pure state. The following demonstrates that a large coherence of the conserved quantity in the initial state of the apparatus, when such a state may be mixed, is a necessary condition for accurate and non-disturbing measurements in the presence of a full conservation law.

Proposition 3.1. *Consider again the setup of Theorem 3.1, and assume that the interaction channel \mathcal{E} also fully conserves $N = N_S \otimes \mathbb{1}_A + \mathbb{1}_S \otimes N_A$. Then for all $x \in \mathcal{X}$ it also holds that*

$$\|[\mathbf{E}(x), N_S] - \Lambda^*([Z(x), N_A])\| \leq 2\|N_S\| \|\epsilon(x)\| + \frac{1}{2}\mathcal{Q}(N_A, \xi)^{\frac{1}{2}}, \quad (14)$$

and for all $y \in \mathcal{Y}$ it also holds that

$$\|[\mathbf{F}(y), N_S] - \mathcal{I}_{\mathcal{X}}^*([\mathbf{F}(y), N_S])\| \leq 2\|N_S\| \|\delta(y)\| + \frac{1}{2}\mathcal{Q}(N_A, \xi)^{\frac{1}{2}}. \quad (15)$$

The proof for Eq. (14) is provided in Appendix (G), and the proof for Eq. (15) is given in Appendix (H). Let us stress that since full conservation implies average conservation, then the inequalities in the above proposition are *complementary* to those given in Theorem 3.1. That is, when the observables are sharp, then zero error and zero disturbance are possible only if $[\mathbf{E}(x), N_S] = \Lambda^*([Z(x), N_A])$ and $[\mathbf{F}(y), N_S] = \mathcal{I}_{\mathcal{X}}^*([\mathbf{F}(y), N_S])$, respectively, independently of the coherence in the apparatus preparation. This is because when the observables are sharp and there is zero error for E or zero disturbance for F, while the upper bounds in Proposition 3.1 may be large, the upper bounds in Theorem 3.1 vanish. Notwithstanding, we see that so long as the apparatus preparation has a large coherence, approximately accurate measurements for E and approximate non-disturbance for F will not be ruled out, even when these observables are sharp.

3.3 The generalised Wigner-Araki-Yanase theorem

We are now ready to give two formulations of the WAY theorem which go beyond existing work in several respects. The first formulation is a direct consequence of the quantitative bounds given above.

Theorem 3.2 (Generalised WAY theorem 1). *Let $\mathcal{M} := (\mathcal{H}_A, \xi, \mathcal{E}, Z)$ be a measurement scheme for an E-instrument \mathcal{I} acting in \mathcal{H}_S , and assume that \mathcal{E} conserves an additive quantity $N = N_S \otimes \mathbb{1}_A + \mathbb{1}_S \otimes N_A$ on average, where $N_S \in \mathcal{L}_s(\mathcal{H}_S)$ and $N_A \in \mathcal{L}_s(\mathcal{H}_A)$. If either \mathcal{I} is repeatable, or the Yanase condition $[Z, N_A] = \mathbf{0}$ is satisfied, then for all $x \in \mathcal{X}$ it holds that*

$$\|[\mathbf{E}(x), N_S]\| \leq 2\|\Gamma_\xi^\mathcal{E}(N^2) - \Gamma_\xi^\mathcal{E}(N)^2\|^{\frac{1}{2}} \|\mathbf{E}(x) - \mathbf{E}(x)^2\|^{\frac{1}{2}}, \quad (16)$$

where $\Gamma_\xi^\mathcal{E}$ is the channel defined in Eq. (3). If \mathcal{E} also fully conserves N , and if either \mathcal{I} is repeatable or the Yanase condition is satisfied, then for all $x \in \mathcal{X}$ it also holds that

$$\|[\mathbf{E}(x), N_S]\| \leq \frac{1}{2}\mathcal{Q}(N_A, \xi)^{\frac{1}{2}}, \quad (17)$$

where $\mathcal{Q}(N_A, \xi)$ is the quantum Fisher information of N_A in ξ as defined in Eq. (13).

Proof. We first prove Eq. (16). By Theorem 3.1, and setting $\epsilon = 0$, if \mathcal{M} is a measurement scheme for E then it must hold that

$$\|[\mathbf{E}(x), N_S] - \Lambda^*([Z(x), N_A])\| \leq 2\|\Gamma_\xi^\mathcal{E}(N^2) - \Gamma_\xi^\mathcal{E}(N)^2\|^{\frac{1}{2}} \|\mathbf{E}(x) - \mathbf{E}(x)^2\|^{\frac{1}{2}} \quad \forall x \in \mathcal{X}.$$

If the Yanase condition is satisfied, then $\Lambda^*([Z(x), N_A]) = \Lambda^*(\mathbf{0}) = \mathbf{0}$, and so we obtain Eq. (16).

Now let Z be an arbitrary pointer observable, but assume that \mathcal{I} is a repeatable E-instrument. Recall that repeatability implies first-kindness, which is a specific instance of non-disturbance. Then by Theorem 3.1, identifying F with E, and setting $\delta = 0$, it must hold that

$$\|[\mathbf{E}(x), N_S] - \mathcal{I}_{\mathcal{X}}^*([\mathbf{E}(x), N_S])\| \leq 2\|\Gamma_\xi^\mathcal{E}(N^2) - \Gamma_\xi^\mathcal{E}(N)^2\|^{\frac{1}{2}} \|\mathbf{E}(x) - \mathbf{E}(x)^2\|^{\frac{1}{2}} \quad \forall x \in \mathcal{X}.$$

By item (vi) of Proposition E.1, if \mathcal{I} is a repeatable measurement of E , then $\mathcal{I}_X^*(E(x)A) = \mathcal{I}_X^*(AE(x))$ for all $A \in \mathcal{L}(\mathcal{H}_S)$. It follows that $\mathcal{I}_X^*([E(x), N_S]) = \mathbf{0}$, and so once again we obtain Eq. (16). Indeed, let us note that if the measurement is repeatable, then for any choice of pointer observable Z it will also hold that $\Lambda^*(Z(x)B) = \Lambda^*(BZ(x))$ for all $B \in \mathcal{L}(\mathcal{H}_A)$, and so $\Lambda^*([Z(x), N_A]) = \mathbf{0}$ even if Z violates the Yanase condition. It follows that we can obtain Eq. (16) under the repeatability assumption directly from the measurability bound. By the same arguments as above, Eq. (17) is obtained from Proposition 3.1. \square

This theorem goes beyond the original WAY theorem (and its descendants) in the setting of bounded conserved quantities in the following respects: it holds for general interaction channels, unsharp target observables, unsharp pointer observables, and mixed apparatus states. It also provides an operationally motivated quantitative bound from which the original theorem can be obtained as a special case: if E is a sharp observable, the upper bound of Eq. (16) vanishes, in which case an additive conservation law together with either repeatability or the Yanase condition necessitates commutation of E with the system part of the conserved quantity. Note that the impossibility statement of the original WAY theorem holds even under the weaker notion of average conservation. This shows that the impossibility of perfect measurements, for sharp observables not commuting with the conserved quantity, holds in much broader contexts than previously assumed. Indeed, such constraint holds even when the measurement is not constrained by ‘‘symmetry’’; recall that while full conservation of N by \mathcal{E} implies that \mathcal{E} is covariant with respect to unitary evolution generated by N (in fact, it is also invariant), it may be the case that \mathcal{E} conserves N on average without being covariant.

Theorem 3.2 does not rule out accurate (under the Yanase condition) or repeatable measurements for unsharp observables not commuting with the conserved quantity, provided an appropriate apparatus preparation: in the special case of a full conservation law, the apparatus preparation must have a large coherence in the conserved quantity. However, this does not imply that coherence allows for accurate or repeatable measurements of *all* unsharp observables. We now present a further generalisation of the WAY theorem, providing additional necessary conditions for perfect measurements that are independent of the apparatus preparation.

Theorem 3.3 (Generalised WAY theorem 2). *Let $\mathcal{M} := (\mathcal{H}_A, \xi, \mathcal{E}, Z)$ be a measurement scheme for an E -instrument \mathcal{I} acting in \mathcal{H}_S , and assume that \mathcal{E} conserves an additive quantity $N = N_S \otimes \mathbf{1}_A + \mathbf{1}_S \otimes N_A$ on average, where $N_S \in \mathcal{L}_s(\mathcal{H}_S)$ and $N_A \in \mathcal{L}_s(\mathcal{H}_A)$. If either \mathcal{I} is a measurement of the first kind, or the Yanase condition $[Z, N_A] = \mathbf{0}$ is satisfied, then for any effect $E(x)$ that has both eigenvalue 1 and 0, it holds that*

$$P[E(x), N_S]P = [E(x), PN_S]P = \mathbf{0}, \quad (18)$$

where $P := P_0(x) + P_1(x)$, with $P_0(x)$ and $P_1(x)$ orthogonal projections onto the eigenvalue-0 and eigenvalue-1 eigenspaces of $E(x)$, respectively.

For a proof, see Appendix (I). The first equality in Eq. (18) follows from the fact that $E(x)P = PE(x) = P_1(x)$, so that $E(x)$ commutes with P , and the second equality states that while the commutator $[E(x), N_S]$ may not vanish entirely, it does vanish when projected onto the subspace $P\mathcal{H}_S$. The above theorem is an even stronger extension of the original WAY theorem, as it relaxes the repeatability condition to that of first-kindness; recall that while a measurement that is repeatable is also of the first kind, repeatability and first-kindness coincide only for sharp observables, and a measurement of an unsharp observable may be of the first kind but not repeatable. Moreover, note that an observable admits a repeatable measurement only if all effects have eigenvalue 1 which, by normalisation, implies that all effects have both eigenvalue 1 and 0. In such a case, Eq. (18) applies to every effect, and P may be interpreted as the projection onto the union of eigenvalue-1 eigenspaces of all the effects of E , i.e., $P = \sum_{x \in \mathcal{X}} P_1(x)$. Finally, note that if E is sharp, then $P = \mathbf{1}_S$, in which case the original WAY theorem is once again recovered.

The condition of an effect $E(x)$ having both eigenvalue 1 and 0 implies that the effect is *definite*, or admits definite values. Specifically, such a condition implies that there exist states ρ for which outcome x can be predicted to obtain with probabilistic certainty, i.e., $\text{tr}[E(x)\rho] = 1$, and that there exist states σ for which outcome x can be predicted to *not* obtain with probabilistic certainty, i.e., $\text{tr}[E(x)\sigma] = 0$. Therefore, a measurement of such an E allows for perfect distinguishability of states ρ and σ ; if outcome x is observed, we know with probabilistic certainty that the system was not prepared in state σ . Conversely, if any outcome $y \neq x$ is observed, we know with probabilistic certainty that the system was not prepared in state ρ . Theorem 3.3 therefore demonstrates that the impossibility part of the WAY theorem—originally pertaining to sharpness—is more properly understood as concerning observables with definite values, even if unsharp. That is to say, if we wish to achieve perfectly accurate (under the Yanase condition) or first-kind measurements of an observable that truly does not commute with the conserved quantity, i.e., such that $[E(x), N_S]$ does not vanish even when projected onto a subspace of \mathcal{H}_S , then not only must such an observable be unsharp, but it must also not admit definiteness.

Theorem 3.3 imposes stronger constraints than Theorem 3.2 and demonstrates that in general, and irrespective of the apparatus preparation, there exist unsharp observables not commuting with the conserved quantity that do not admit a repeatable or first-kind measurement, and which cannot be accurately measured if the Yanase condition holds. To illustrate this, let us introduce the following model. Consider a system $\mathcal{H}_S \simeq \mathbb{C}^3$ with the orthonormal basis $\{|-1\rangle, |0\rangle, |1\rangle\}$, and the conserved quantity $N_S = \sum_n n|n\rangle\langle n| \equiv |1\rangle\langle 1| - |-1\rangle\langle -1|$. Consider also the class of binary observables $E_\lambda := \{E_\lambda(+), E_\lambda(-)\}$ acting in \mathcal{H}_S , defined by

$$E_\lambda(\pm) := \lambda|\pm\rangle\langle\pm| + (1-\lambda)|\mp\rangle\langle\mp| + \frac{1}{2}|0\rangle\langle 0|, \quad (19)$$

where $1/2 < \lambda \leq 1$ and $|\pm\rangle := \frac{1}{\sqrt{2}}(|1\rangle \pm |-1\rangle)$. In the absence of any constraints, all observables in this class admit first-kind measurements. For example, since E_λ is commutative, then the corresponding Lüders instrument is a first-kind (but not repeatable) measurement of E_λ . On the other hand, E_λ admits a repeatable measurement if and only if $\lambda = 1$, in which case an instrument with operations $\mathcal{I}_\pm(\cdot) = \text{tr}[E_\lambda(\pm)\cdot]|\pm\rangle\langle\pm|$ is a repeatable measurement of E_λ . Since E_λ is unsharp, even when $\lambda = 1$, then Theorem 3.2 does not rule out accurate or repeatable measurements for such an observable, provided an appropriate apparatus preparation. But now note that when $\lambda = 1$, both effects have eigenvalue 1 and 0, and we have $P_1(\pm) = |\pm\rangle\langle\pm|$ and $P_0(\pm) = |\mp\rangle\langle\mp|$, and so $P = P_0(\pm) + P_1(\pm) = |1\rangle\langle 1| + |-1\rangle\langle -1|$. It is easily verified that in such a case,

$$[E_\lambda(\pm), PN_S P] = [E_\lambda(\pm), N_S] = |\mp 1\rangle\langle\pm 1| - |\pm 1\rangle\langle\mp 1| \neq \mathbf{0}.$$

By Theorem 3.3, it follows that when $\lambda = 1$, a measurement of E_λ that is constrained by an average conservation law cannot be repeatable or even first-kind, and must violate the Yanase condition. However, the effects of E_λ when $\lambda < 1$ do not commute with the conserved quantity, and have neither eigenvalue 1 nor eigenvalue 0. In such a case, Theorem 3.3 does not rule out accurate or first-kind (but not repeatable) measurements.

3.4 The Wigner-Araki-Yanase theorem without the Yanase condition

Traditionally, the Yanase condition is justified by applying the repeatability part of the WAY theorem to the pointer observable; if the pointer observable is sharp, and we consider its measurement as being implemented by a conservative interaction between one measuring apparatus and another, then the pointer observable will admit a repeatable measurement only if it commutes with the conserved quantity. Repeatability of the measurement of the pointer observable is deemed a natural requirement for the possibility of measurement, since an experimenter should be able to confirm the measurement outcome by repeated observations of the apparatus: there must be a stable record of the measurement outcomes. However, such an argument suffers from two drawbacks. Firstly, it applies only to sharp pointer observables. Secondly, it runs into the problem of infinite regress, since we have now shifted the role of the ultimate pointer observable from the first apparatus to the second; repeatability of the first pointer observable can be abandoned if the second admits a repeatable measurement, in which case the experimenter may continue to verify the measurement outcomes. In Appendix (J), we show that the measurability part of the WAY theorem—Theorem 3.2 and Theorem 3.3—can be justified without an appeal to the Yanase condition, but rather by imposing a conservation law on the total measurement process, i.e., including pointer objectification. Such conservation is shown to give rise to the so-called “weak” Yanase condition $[Z^T, N] = \mathbf{0}$, where $Z^T(x) := \mathcal{E}^*(\mathbf{1}_S \otimes Z(x))$ is the “Heisenberg-evolved” pointer observable [32]. Subsequently, it is shown that if the weak Yanase condition is satisfied, then Eq. (16) (with $\|\Gamma_\xi^\mathcal{E}(N^2) - \Gamma_\xi^\mathcal{E}(N)^2\|$ replaced by $\text{Var}(N_A, \xi)$), Eq. (17), and Eq. (18) will hold.

4 Fixed points and non-disturbing measurements in the presence of conservation laws

In Sec. 3 we provided general quantitative bounds for measurement disturbance, when observables E and F are measured in succession and when the first measurement is subject to a conservation law. As we saw, these bounds generally do not prohibit non-disturbance for an unsharp F that does not commute with the conserved quantity. In this section, we provide tighter restrictions on the possibility of non-disturbance that depend on the structure of the fixed-point set $\mathcal{F}(\mathcal{I}_\mathcal{X}^*)$, which intimately depends on the properties of observable E and the states that are left invariant by the E -channel $\mathcal{I}_\mathcal{X}$. We first consider the case where $\mathcal{F}(\mathcal{I}_\mathcal{X}^*)$ is a von Neumann algebra, which is guaranteed to be the case when $\mathcal{F}(\mathcal{I}_\mathcal{X})$ contains a faithful state. Next, we relax the faithfulness condition on the states in $\mathcal{F}(\mathcal{I}_\mathcal{X})$, and obtain similar restrictions for non-disturbance in the finite-dimensional setting. Finally, we show that in the finite-dimensional case, the first-kindness statement of our generalisation of WAY in Theorem 3.3 can be extended to a quantitative bound.

4.1 Non-disturbance and von Neumann algebras

By Theorem 3.1, when a sharp observable F is not disturbed, then the upper bound of Eq. (12) vanishes, implying that non-disturbance is possible only if $[F(y), N_s] \in \mathcal{F}(\mathcal{I}_\chi^*)$. But in Appendix (H) we provide a tighter upper bound than that of Eq. (12), which vanishes if both $F \in \mathcal{F}(\mathcal{I}_\chi^*)$ and $F^2 := \{F(y)^2 : y \in \mathcal{Y}\} \subset \mathcal{F}(\mathcal{I}_\chi^*)$ hold, in which case non-disturbance will be possible only if $[F(y), N_s] \in \mathcal{F}(\mathcal{I}_\chi^*)$. While non-disturbance of F trivially implies that $F^2 \subset \mathcal{F}(\mathcal{I}_\chi^*)$ when F is sharp, the implication $F \in \mathcal{F}(\mathcal{I}_\chi^*) \implies F^2 \subset \mathcal{F}(\mathcal{I}_\chi^*)$ holds for all observables whenever $\mathcal{F}(\mathcal{I}_\chi^*)$ is a von Neumann algebra, which is guaranteed to be the case whenever $\mathcal{F}(\mathcal{I}_\chi)$ contains at least one faithful state. We now show that in the presence of conservation laws, if the fixed-point set of the measurement channel \mathcal{I}_χ^* is a von Neumann algebra, there are strong constraints imposed on the possibility of non-disturbance.

Theorem 4.1. *Let $E := \{E(x) : x \in \mathcal{X}\}$ and $F := \{F(y) : y \in \mathcal{Y}\}$ be observables acting in \mathcal{H}_s . Let $\mathcal{M} := (\mathcal{H}_A, \xi, \mathcal{E}, \mathcal{Z})$ be a measurement scheme for an E -instrument \mathcal{I} , and assume that \mathcal{E} conserves an additive quantity $N = N_s \otimes \mathbb{1}_A + \mathbb{1}_s \otimes N_A$ on average, where $N_s \in \mathcal{L}_s(\mathcal{H}_s)$ and $N_A \in \mathcal{L}_s(\mathcal{H}_A)$. If $\mathcal{F}(\mathcal{I}_\chi^*)$ is a von Neumann algebra, then the following hold:*

- (i) $F \in \mathcal{F}(\mathcal{I}_\chi^*)$ only if F commutes with E , with $\Delta N_s := \mathcal{I}_\chi^*(N_s) - N_s$, and with $\{[E(x), N_s] : x \in \mathcal{X}\}$.
- (ii) \mathcal{I} is a measurement of the first kind only if E is commutative and commutes with N_s .
- (iii) \mathcal{I} is a repeatable measurement only if E is sharp and commutes with N_s .

We note that the condition $[E, F] = \mathbb{O}$ in item (i) is independent of conservation, and was already shown in Ref. [4]. The proof of the above theorem is given in Appendix (L) (Theorem L.1), and here is a rough sketch for item (i). If $\mathcal{F}(\mathcal{I}_\chi^*)$ is an algebra, then $\mathcal{F}(\mathcal{I}_\chi^*) \subset E'$, and so non-disturbance is possible only if $[E, F] = \mathbb{O}$. But if the measurement obeys a conservation law, and $F \in \mathcal{F}(\mathcal{I}_\chi^*)$, then additionally it holds that $[F(y), N_s] \in \mathcal{F}(\mathcal{I}_\chi^*)$, which by the multiplicability theorem implies that $[F(y), \Delta N_s] = \mathbb{O}$. Finally, by $[F(y), N_s] \in \mathcal{F}(\mathcal{I}_\chi^*) \subset E'$ and $[E, F] = \mathbb{O}$, it holds that $[F(y), [E(x), N_s]] = [E(x), [F(y), N_s]] = \mathbb{O}$.

Assuming that F commutes with E , the additional necessary condition for non-disturbance, $[F(y), [E(x), N_s]] = \mathbb{O}$, is guaranteed to hold if either E commutes with N_s , or if F commutes with N_s . Note that since $\mathcal{F}(\mathcal{I}_\chi^*) \subset E'$ when the fixed-point set of the measurement channel is an algebra, whenever E does not commute with N_s then F must also commute with $\Delta N_s \neq \mathbb{O}$, which is not in general guaranteed by commutation of F with N_s . Of course, unless $E = F$, non-disturbance may be possible even if neither E nor F commute with N_s . But by items (ii) and (iii) of the above theorem, when $E = F$, i.e., when \mathcal{I} is a first-kind or repeatable measurement of E , then non-disturbance is possible only if E commutes with N_s . Indeed, we see that when the fixed-point set of the measurement channel is an algebra, then the constraints on repeatability and first-kindness are much stronger than in the more general case as given by Theorem 3.2 and Theorem 3.3. We may therefore strengthen the necessary conditions for repeatability and first-kindness that are given by the WAY theorem with the following: in the presence of a conservation law, an unsharp and possibly non-commutative observable E not commuting with the conserved quantity admits a repeatable or first-kind measurement only if the E -channel \mathcal{I}_χ perturbs all faithful states.

Let us now consider some interesting consequences of the above theorem. As shown by Proposition 6 in Ref. [4], when the system is a qubit, i.e., $\dim(\mathcal{H}_s) = 2$, then $\mathcal{F}(\mathcal{I}_\chi^*)$ is an algebra for any instrument \mathcal{I} . See also Corollary M.1 in Appendix (M.2). It follows that for qubits, the implications of Theorem 4.1 will hold in general. Now let us assume that E is a binary observable with the effects

$$E(\pm) = \frac{1}{2}(\mathbb{1}_s \pm \lambda \sigma_1),$$

where $\lambda \in (0, 1]$ and $\sigma_1, \sigma_2, \sigma_3$ are the Pauli operators. This observable is sharp when $\lambda = 1$, and is unsharp when $\lambda < 1$. Since binary observables are commutative, then item (ii) of Theorem 4.1 will permit a first-kind measurement of E so long as the conserved quantity commutes with σ_1 . On the other hand, by item (iii) repeatability will be allowed only if $\lambda = 1$ also holds. This is not so surprising since repeatability is permitted only when all effects have eigenvalue 1, with such condition being satisfied for qubit observables only when the observable is sharp. Now let us assume that the conserved quantity is $N_s = \sigma_3$, so that it does not commute with E , which implies that repeatability and first-kindness will be ruled out. But can a measurement of E not disturb some other observable? Note that $[E(\pm), \sigma_3] = \mp \lambda i \sigma_2$. By item (i), a non-trivial observable F will be non-disturbed only if it commutes with σ_1 and with σ_2 , which is clearly impossible. Indeed, for qubits in the presence of conservation, non-disturbing measurements are permitted only if $[E, F] = [E, N_s] = [F, N_s] = \mathbb{O}$.

Now let us consider the case where an observable E is measured by the Lüders instrument \mathcal{I}^L defined in Eq. (1). If E is commutative, then it holds that $\mathcal{F}(\mathcal{I}_\chi^{L*}) = E'$ is a von Neumann algebra (as the commutant of a self-adjoint subset of \mathcal{H}_s), even in infinite dimensions [76, 77]. But recall that the Lüders instrument is a first-kind

measurement if it is compatible with a commutative observable. By item (ii) of Theorem 4.1 it follows that in the presence of a conservation law, a commutative E admits a Lüders instrument only if E commutes with the system part of the conserved quantity. Now let us consider an observable E that may be non-commutative, but commutes with N_S . Note that in this case, unless $\dim(\mathcal{H}_S) < \infty$, then $\mathcal{F}(\mathcal{I}_{\mathcal{X}}^{L*})$ is not necessarily an algebra, since in infinite dimensions there exist non-commutative observables E such that $\mathcal{F}(\mathcal{I}_{\mathcal{X}}^{L*}) \not\subset E'$ [78, 79]. But since $E' \subset \mathcal{F}(\mathcal{I}_{\mathcal{X}}^{L*})$ always holds, then $[E, N_S] = \mathbf{0} \implies N_S^k \in \mathcal{F}(\mathcal{I}_{\mathcal{X}}^{L*})$ for all $k \in \mathbb{N}$. That is, the Lüders measurement of E commuting with N_S will fully conserve N_S . Recall that in such a case, non-disturbance will not be ruled out for any observable F that commutes with E (see Corollary H.2). Indeed, it will hold that $\Delta N_S = \mathbf{0}$, and so any F will trivially commute with ΔN_S . But does an observable E commuting with N_S always admit a Lüders measurement in the presence of a conservation law? We shall now show that in the presence of a full conservation law, and where the apparatus part of the conserved quantity is highly non-degenerate, such measurements will require a large coherence in the apparatus preparation. This surprising observation can be seen as a “converse” WAY theorem.

Proposition 4.1. *Let $\mathcal{M} := (\mathcal{H}_A, \xi, \mathcal{E}, \mathcal{Z})$ be a measurement scheme for an instrument \mathcal{I} acting in \mathcal{H}_S . Assume that \mathcal{E} fully conserves an additive quantity $N = N_S \otimes \mathbb{1}_A + \mathbb{1}_S \otimes N_A$, where $N_S \in \mathcal{L}_s(\mathcal{H}_S)$ and $N_A \in \mathcal{L}_s(\mathcal{H}_A)$, and that $\mathcal{I}_{\mathcal{X}}$ fully conserves N_S . Define the subspace of \mathcal{H}_A that is involved during the measurement process as*

$$\mathcal{H}_A(\text{meas}) := \bigcup_{\rho \in \mathcal{S}(\mathcal{H}_S)} \text{supp}(\Lambda(\rho)) \cup \text{supp}(\xi) \subseteq \mathcal{H}_A,$$

where Λ is the conjugate channel to $\mathcal{I}_{\mathcal{X}}$ defined in Eq. (5). Then either $\mathcal{H}_A(\text{meas})$ is contained within a single degenerate eigenspace of N_A , or $\text{Var}(N_A, \xi) := \text{tr}[N_A^2 \xi] - \text{tr}[N_A \xi]^2$ must be large. Additionally, if \mathcal{I} is an extremal instrument, and if $\mathcal{H}_A(\text{meas})$ is not contained within a single degenerate eigenspace of N_A , then the quantum Fisher information $\mathcal{Q}(N_A, \xi)$ as defined in Eq. (13) must be large.

Note that even if \mathcal{I} is not a Lüders instrument, by item (iii) of Lemma C.1 it holds that if \mathcal{I} is compatible with observable E then $\mathcal{I}_{\mathcal{X}}$ fully conserves N_S only if $[E, N_S] = \mathbf{0}$. Moreover, let us note that the Lüders E -instrument is extremal whenever the effects of E are linearly independent [80]. The proof of the above proposition is provided in Appendix (K), and here we present a rough sketch. In the case that \mathcal{E} fully conserves N and $\mathcal{I}_{\mathcal{X}}$ fully conserves N_S then both the expected value and the variance of N_A must not change as a result of the measurement interaction. That is, $\text{tr}[N_A \Lambda(\rho)] = \text{tr}[N_A \xi]$ and $\text{Var}(N_A, \Lambda(\rho)) = \text{Var}(N_A, \xi)$ for all $\rho \in \mathcal{S}(\mathcal{H}_S)$. It follows that if ξ is an eigenstate of N_A , i.e., if ξ has support only in a single degenerate eigenspace of N_A , then $\Lambda(\rho)$ must live in the same eigenspace for all ρ . That is, N_A must be “effectively” fully degenerate, in the sense that $\mathcal{H}_A(\text{meas})$ must be contained within a single degenerate eigenspace of N_A . This generalises an observation made in Ref. [81], which held only in the case of Lüders measurements of sharp observables, implemented by normal measurement schemes satisfying the Yanase condition. But in many physically relevant situations N_A will not be (effectively) fully degenerate—for example, the apparatus may be a system with a conserved quantity N_A that is completely non-degenerate. In such cases, when the interaction between system and apparatus obeys a full conservation law, an instrument that fully conserves N_S can be implemented only if the apparatus preparation is not an eigenstate of N_A , which implies that $\text{Var}(N_A, \xi)$ must be large. Finally, if the instrument is extremal, and N_A is not (effectively) fully degenerate, then for every pure state decomposition $\xi = \sum_i q_i P_{\phi_i}$, the uncertainty of N_A in ϕ_i must be large, which implies that the apparatus preparation must have a large coherence as quantified by the quantum Fisher information.

4.2 Non-disturbance and operator spaces

Due to the Schauder–Tychonoff fixed point theorem [82], all E -channels $\mathcal{I}_{\mathcal{X}}$ have at least one fixed state. However, it may be that none of these are faithful. In such a case, the fixed-point set of the dual channel $\mathcal{I}_{\mathcal{X}}^*$ is not necessarily a von Neumann algebra, but rather forms an operator space [83]. This setting has been much less investigated, and in Appendix (M.1) we provide some novel analysis of the structure of such fixed-point sets. While the discussion thus far has been applicable for infinite-dimensional systems—except in some examples—in this section we shall always assume that $\dim(\mathcal{H}_S) < \infty$.

We define the minimal support projection P on the fixed-point set $\mathcal{F}(\mathcal{I}_{\mathcal{X}})$ as

$$P = \min\{Q : Q \text{ is a projection, } \rho = Q\rho Q \forall \rho \in \mathcal{F}(\mathcal{I}_{\mathcal{X}})\}. \quad (20)$$

In other words, for all projections Q and fixed states $\rho \in \mathcal{F}(\mathcal{I}_{\mathcal{X}})$ such that $\rho = Q\rho Q$, it holds that $Q \geq P$. Note that $P = \mathbb{1}_S$ if and only if $\mathcal{F}(\mathcal{I}_{\mathcal{X}})$ contains a faithful state, in which case $\mathcal{F}(\mathcal{I}_{\mathcal{X}}^*)$ is an algebra, so that we recover the results of Theorem 4.1. We now provide a generalisation of this result which accounts for situations where P may be smaller than the identity, i.e., where the E -channel $\mathcal{I}_{\mathcal{X}}$ may perturb all faithful states. Here,

we define $PEP := \{PE(x)P : x \in \mathcal{X}\}$ and $PFP := \{PF(y)P : y \in \mathcal{Y}\}$ as restrictions of observables E and F in $P\mathcal{H}_S$.

Theorem 4.2. *Let $E := \{E(x) : x \in \mathcal{X}\}$ and $F := \{F(y) : y \in \mathcal{Y}\}$ be observables acting in \mathcal{H}_S . Let $\mathcal{M} := (\mathcal{H}_A, \xi, \mathcal{E}, \mathcal{Z})$ be a measurement scheme for an E -instrument \mathcal{I} , and assume that \mathcal{E} conserves an additive quantity $N = N_S \otimes \mathbb{1}_A + \mathbb{1}_S \otimes N_A$ on average, where $N_S \in \mathcal{L}_s(\mathcal{H}_S)$ and $N_A \in \mathcal{L}_s(\mathcal{H}_A)$. If P is the minimal support projection on $\mathcal{F}(\mathcal{I}_\mathcal{X})$, then the following hold:*

- (i) $F \subset \mathcal{F}(\mathcal{I}_\mathcal{X}^*)$ only if PFP commutes with PEP , with $P\Delta N_S P := P\mathcal{I}_\mathcal{X}^*(N_S)P - PN_S P$, and with $\{[PE(x)P, PN_S P] : x \in \mathcal{X}\}$.
- (ii) \mathcal{I} is a measurement of the first kind only if PEP is commutative and commutes with $PN_S P$.
- (iii) \mathcal{I} is repeatable only if PEP is sharp and commutes with $PN_S P$.

The proof is provided in Appendix (M.2) (Theorem M.1) and it follows from similar arguments as those used in Theorem 4.1. That is, by noting that there exists a faithful fixed state in the subspace $P\mathcal{H}_S$, we observe that the projection of the fixed-point set $\mathcal{F}(\mathcal{I}_\mathcal{X}^*)$ onto the subspace $P\mathcal{H}_S$ is a von Neumann algebra.

As a simple example, let us consider the case where E is measured by a *nuclear* instrument. The operations of a nuclear instrument \mathcal{I} are written as $\mathcal{I}_x(\cdot) = \text{tr}[E(x)\cdot]\sigma_x$, where $\{\sigma_x\}$ is a family of states. It is simple to verify that in such a case, P is the minimal projection on $\cup_x \text{supp}(\sigma_x)$. Additionally, if E is a norm-1 observable, and for every x the support of σ_x is contained within the eigenvalue-1 eigenspace of $E(x)$, then such an instrument will be a repeatable measurement of E . Every observable admits a nuclear instrument and, as shown in Corollary 1 of Ref. [4], every instrument compatible with a rank-1 observable is nuclear. Now assume that \mathcal{I} does not disturb some observable F . Since the dual E -channel may be written as $\mathcal{I}_\mathcal{X}^*(\cdot) = \sum_x \text{tr}[\cdot\sigma_x]E(x)$, we observe that $F \subset \mathcal{F}(\mathcal{I}_\mathcal{X}^*)$ implies that $F(y) = \sum_x p(y|x)E(x)$. That is, non-disturbance is possible only if F is a classical post-processing of E . But note that unless E is commutative, this does not generally imply that F must commute with E . However, item (i) of Theorem 4.2 states that PFP must commute with PEP and with $P\Delta N_S P$. Given that $P\Delta N_S P = P\mathcal{I}_\mathcal{X}^*(N_S)P - PN_S P = \sum_x \text{tr}[N_S\sigma_x]PE(x)P - PN_S P$, then $[PFP, PEP] = \mathbb{O}$ implies that the commutation of PFP with $P\Delta N_S P$ simplifies to $[PFP, PN_S P] = \mathbb{O}$. Indeed, in the case where $\cup_x \text{supp}(\sigma_x) = \mathcal{H}_S$ so that $P = \mathbb{1}_S$, it follows that only observables F that commute with both the measured observable E , and with the system part of the conserved quantity N_S , will be non-disturbed.

While the implications of the above theorem depend on the minimal support projection P on the fixed states of the measurement channel $\mathcal{I}_\mathcal{X}$, and hence on the specific measurement implementation, we may use the structure of the fixed-point set to obtain necessary conditions for first-kindness that depend only on the measured observable. In Appendix (M.3), we provide some necessary conditions for non-disturbance that are independent of conservation laws, showing that non-disturbance is intimately related to distinguishability. In particular, we show that if \mathcal{I} is a first-kind measurement of E , then this observable must be a classical post-processing of a norm-1 observable $G \subset \mathcal{F}(\mathcal{I}_\mathcal{X}^*)$, and that there exists a family of states $\{\rho_z\}$ that are perfectly distinguishable by a measurement of G such that $\{\mathcal{I}_\mathcal{X}(\rho_z)\}$ remain perfectly distinguishable. Next, we use these results to obtain a quantitative form of the WAY theorem for first-kindness, presented below.

Theorem 4.3. *Consider a measurement scheme $\mathcal{M} := (\mathcal{H}_A, \xi, \mathcal{E}, \mathcal{Z})$ for a nontrivial observable E with the instrument \mathcal{I} acting in \mathcal{H}_S . Assume that \mathcal{I} is a measurement of the first kind, and that \mathcal{E} conserves an additive quantity $N = N_S \otimes \mathbb{1}_A + \mathbb{1}_S \otimes N_A$ on average, where $N_S \in \mathcal{L}_s(\mathcal{H}_S)$ and $N_A \in \mathcal{L}_s(\mathcal{H}_A)$. For each outcome x associated with a non-trivial effect $E(x)$, let $\mathcal{K}_{\max}(x)$ and $\mathcal{K}_{\min}(x)$ be subspaces of \mathcal{H}_S defined by*

$$\mathcal{K}_{\max}(x) := \{\psi \in \mathcal{H}_S : E(x)\psi = \|E(x)\|\psi\}, \quad \mathcal{K}_{\min}(x) := \{\phi \in \mathcal{H}_S : (\mathbb{1}_S - E(x))\phi = \|\mathbb{1}_S - E(x)\|\phi\}.$$

$\mathcal{K}_{\max}(x)$ and $\mathcal{K}_{\min}(x)$ are orthogonal, and for all unit vectors $\psi \in \mathcal{K}_{\max}(x)$ and $\phi \in \mathcal{K}_{\min}(x)$, it holds that

$$|\langle \psi | N_S \phi \rangle| \leq \|N_S\| \left(\|E(x)\|^{\frac{1}{2}} (1 - \|\mathbb{1}_S - E(x)\|)^{\frac{1}{2}} + (1 - \|E(x)\|)^{\frac{1}{2}} \|\mathbb{1}_S - E(x)\|^{\frac{1}{2}} \right). \quad (21)$$

For a proof, see Appendix (M.4) (Theorem M.3). We note that if E commutes with N_S , then Theorem 4.3 imposes no restrictions on first-kindness. This is because so long as $E(x)$ is non-trivial and commutes with N_S , then the lower bound of Eq. (21) vanishes. On the other hand, if for any non-trivial $E(x)$ not commuting with N_S it holds that $\langle \psi | N_S \phi \rangle \neq 0$ for some $\psi \in \mathcal{K}_{\max}(x)$ and $\phi \in \mathcal{K}_{\min}(x)$, then E admits a first-kind measurement only if either (i) $\|E(x)\| < 1$ or (ii) $\|\mathbb{1}_S - E(x)\| < 1$. This is so because if both (i) and (ii) are violated, i.e., if $\|E(x)\| = \|\mathbb{1}_S - E(x)\| = 1$, then the upper bound of Eq. (21) vanishes. It follows that, in such a case, $E(x)$ cannot have both 1 and 0 as an eigenvalue: such an observable must be highly unsharp. Indeed, we may also recover the first-kindness statement of Theorem 3.3 directly from Theorem 4.3, which is presented as Corollary M.3 in Appendix (M.4). This raises an interesting question: will Eq. (21) also hold if we abandon the requirement of first-kindness, and instead assume that the measurement satisfies the Yanase condition? This question is

beyond the scope of the present paper, but the answer may highlight to what extent the necessary conditions for measurability and non-disturbance will continue to satisfy the “symmetry” witnessed so far in WAY-type theorems.

To demonstrate that Theorem 4.3 provides much stronger constraints than Theorem 3.3, let us consider again the simple model of a binary observable $E_\lambda := \{E_\lambda(+), E_\lambda(-)\}$ acting in $\mathcal{H}_S \simeq \mathbb{C}^3$ introduced surrounding Eq. (19). Recall that Theorem 3.3 did not rule out first-kind measurements of E_λ for any $1/2 < \lambda < 1$. But now note that $\mathcal{K}_{\max}(\pm) = \text{span}\{|\pm\rangle\}$, $\mathcal{K}_{\min}(\pm) = \text{span}\{|\mp\rangle\}$, $\|E_\lambda(\pm)\| = \|\mathbb{1}_S - E_\lambda(\pm)\| = \lambda$, $\|N_S\| = 1$, and $|\langle \pm | N_S | \mp \rangle| = 1$. By Theorem 4.3, it follows that such an observable admits a first-kind measurement only if

$$1 \leq 2\sqrt{\lambda(1-\lambda)},$$

which cannot be satisfied for any $1/2 < \lambda < 1$; indeed, the above inequality is satisfied only if $\lambda = 1/2$, in which case $E_\lambda(\pm) = \mathbb{1}_S/2$ are trivial effects.

5 Conclusions

We have provided a number of general and operational bounds which capture measurement error and disturbance, with emphasis on the setting in which there is a conservation law—both “full” conservation, and the weaker notion of “average” conservation. We obtained new, quantitative versions of the WAY theorem, which generalise previous work in several respects, going beyond normal measurement schemes, and not assuming that the observable to be measured is sharp. The work presented surrounding the WAY theorem was also studied in the novel setting of sequential measurements for general pairs of observables, and the quantitative bounds were further refined by the analysis of the fixed point structure of the measurement channel in settings which have received scant attention.

We saw that the large apparatus coherence played a key role for measurability and non-disturbance in the presence of a full conservation law, pointing to the requirement of “large” apparatus. This points further to possible deep connections between the WAY theorem and the rapidly developing theory of quantum reference frames, analysed so far only when the conserved quantity has a conjugate phase [42, 41]. While necessary, however, the large apparatus coherence was shown to not be sufficient for good measurements; we saw that conservation laws impose strict constraints on the error or disturbance for unsharp observables that admit definite values.

Our work suffers from the drawback that many physically arising conserved quantities are unbounded. Very recently, the measurability part of the WAY theorem for sharp target observables was proven in the setting of unbounded conserved quantities, where the conservation law is stated as the invariance of the unitary group generated by the conserved quantity under the action of the measurement interaction [84]. The measurability question for unsharp target observables, as well as the question of disturbance, should also be systematically studied when the conserved quantity is unbounded. This is a technically challenging endeavour and we save it for future work.

Acknowledgments

M.H.M. acknowledges funding from the European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 801505, as well as from the Slovak Academy of Sciences under MoRePro project OPEQ (19MRP0027). T.M. acknowledges financial support from JSPS KAKENHI Grant No.JP20K03732.

A Properties of operations

Operations allow for the construction of an “operator-valued inner product”, which will be frequently used in this paper. For an operation $\Phi^* : \mathcal{L}(\mathcal{K}) \rightarrow \mathcal{L}(\mathcal{H})$, we define the sesquilinear mapping $\langle\langle \cdot | \cdot \rangle\rangle : \mathcal{L}(\mathcal{K}) \times \mathcal{L}(\mathcal{K}) \rightarrow \mathcal{L}(\mathcal{H})$ by

$$\langle\langle A | B \rangle\rangle := \Phi^*(A^*B) - \Phi^*(A^*)\Phi^*(B), \quad (22)$$

to hold for all $A, B \in \mathcal{L}(\mathcal{K})$. The following lemma shows that such a map mimics several important properties of an inner product.

Lemma A.1. For all $A, B, C \in \mathcal{L}(\mathcal{K})$, the sesquilinear mapping defined in Eq. (22) satisfies: (i) $\langle\langle A|B+\lambda C\rangle\rangle = \langle\langle A|B\rangle\rangle + \lambda\langle\langle A|C\rangle\rangle$ for all $\lambda \in \mathbb{C}$; (ii) $\langle\langle A|B\rangle\rangle = \langle\langle B|A\rangle\rangle^*$; (iii) $\langle\langle A|A\rangle\rangle \geq \mathbf{0}$; and (iv) the Cauchy-Schwarz inequality

$$\langle\langle A|B\rangle\rangle\langle\langle B|A\rangle\rangle \leq \|\langle\langle B|B\rangle\rangle\|\langle\langle A|A\rangle\rangle.$$

Proof. (i) trivially follows from linearity of operations, while (ii) follows from the fact that an operation preserves the involution, i.e., $\Phi^*(A)^* = \Phi^*(A^*)$. (iii) follows from Kadison's inequality, or the two-positivity of CP maps [85, 86]. To show this, note that by Stinespring's dilation theorem [87] we may write $\Phi^*(A) = V^*(A \otimes \mathbf{1}_{\mathcal{K}'})V$, where $V : \mathcal{H} \rightarrow \mathcal{K} \otimes \mathcal{K}'$ is a linear operator. Since Φ^* is sub-unital, it must hold that $\Phi^*(\mathbf{1}_{\mathcal{K}}) = V^*(\mathbf{1}_{\mathcal{K}} \otimes \mathbf{1}_{\mathcal{K}'})V \equiv V^*V \leq \mathbf{1}_{\mathcal{H}}$, with equality if Φ^* is a channel, in which case V is an isometry. By the C* identity we therefore have $\|VV^*\| = \|V^*V\| \leq 1$, which implies that $\mathbf{0} \leq VV^* \leq \mathbf{1}_{\mathcal{K}} \otimes \mathbf{1}_{\mathcal{K}'}$. By Eq. (22) we may therefore write

$$\langle\langle A|B\rangle\rangle = V^*(A^* \otimes \mathbf{1}_{\mathcal{K}'})\pi^*\pi(B \otimes \mathbf{1}_{\mathcal{K}'})V, \quad (23)$$

where $\pi = \pi^* := \sqrt{\mathbf{1}_{\mathcal{K}} \otimes \mathbf{1}_{\mathcal{K}'} - VV^*}$. That $\langle\langle A|A\rangle\rangle \geq \mathbf{0}$ trivially follows.

Finally, we prove the Cauchy-Schwarz inequality which, for the case of channels, was proven by Janssens in Lemma 1 of Ref. [53]. The proof for the case of general operations is identical; by Eq. (23) we may write

$$\begin{aligned} \langle\langle A|B\rangle\rangle\langle\langle B|A\rangle\rangle &= V^*(A^* \otimes \mathbf{1}_{\mathcal{K}'})\pi^*\pi(B \otimes \mathbf{1}_{\mathcal{K}'})VV^*(B^* \otimes \mathbf{1}_{\mathcal{K}'})\pi^*\pi(A \otimes \mathbf{1}_{\mathcal{K}'})V \\ &\leq \|\pi(B \otimes \mathbf{1}_{\mathcal{K}'})VV^*(B^* \otimes \mathbf{1}_{\mathcal{K}'})\pi^*\|V^*(A^* \otimes \mathbf{1}_{\mathcal{K}'})\pi^*\pi(A \otimes \mathbf{1}_{\mathcal{K}'})V \\ &= \|V^*(B^* \otimes \mathbf{1}_{\mathcal{K}'})\pi^*\pi(B \otimes \mathbf{1}_{\mathcal{K}'})V\|V^*(A^* \otimes \mathbf{1}_{\mathcal{K}'})\pi^*\pi(A \otimes \mathbf{1}_{\mathcal{K}'})V \\ &= \|\langle\langle B|B\rangle\rangle\|\langle\langle A|A\rangle\rangle. \end{aligned}$$

In the second line we have used the fact that for any self-adjoint operator $A \in \mathcal{L}_s(\mathcal{H})$, it holds that $B^*AB \leq \|A\|B^*B$ for all $B \in \mathcal{L}(\mathcal{H})$, while in the third line we have used the C* identity $\|AA^*\| = \|A^*A\|$ for all $A \in \mathcal{L}(\mathcal{H})$. \square

Note that the sesquilinear mapping in Eq. (22) does not satisfy the positive definiteness property in general, that is, $\langle\langle A|A\rangle\rangle = \mathbf{0}$ does not imply $A = \mathbf{0}$. This plays an important role in the multiplicability theorem [86], which can be seen as a consequence of Lemma A.1:

Corollary A.1. Let $\Phi^* : \mathcal{L}(\mathcal{K}) \rightarrow \mathcal{L}(\mathcal{H})$ be an operation, and consider an operator $B \in \mathcal{L}(\mathcal{K})$. The following hold:

- (i) If $\Phi^*(B^*B) = \Phi^*(B^*)\Phi^*(B)$, then $\Phi^*(AB) = \Phi^*(A)\Phi^*(B)$ for all $A \in \mathcal{L}(\mathcal{K})$.
- (ii) If $\Phi^*(BB^*) = \Phi^*(B)\Phi^*(B^*)$, then $\Phi^*(BA) = \Phi^*(B)\Phi^*(A)$ for all $A \in \mathcal{L}(\mathcal{K})$.

Proof. Let us first prove (i). If $\langle\langle B|B\rangle\rangle = \Phi^*(B^*B) - \Phi^*(B^*)\Phi^*(B) = \mathbf{0}$, then $\|\langle\langle B|B\rangle\rangle\| = 0$. Therefore, by Lemma A.1 we have for all $A \in \mathcal{L}(\mathcal{K})$ the following:

$$\mathbf{0} \leq \langle\langle A^*|B\rangle\rangle\langle\langle A^*|B\rangle\rangle^* = \langle\langle A^*|B\rangle\rangle\langle\langle B|A^*\rangle\rangle \leq \mathbf{0}.$$

This implies that $\langle\langle A^*|B\rangle\rangle = \Phi^*(AB) - \Phi^*(A)\Phi^*(B) = \mathbf{0}$. Similarly for (ii), $\langle\langle B^*|B^*\rangle\rangle = \Phi^*(BB^*) - \Phi^*(B)\Phi^*(B^*) = \mathbf{0}$ implies that for all $A \in \mathcal{L}(\mathcal{K})$ we have

$$\mathbf{0} \leq \langle\langle B^*|A\rangle\rangle^*\langle\langle B^*|A\rangle\rangle = \langle\langle A|B^*\rangle\rangle\langle\langle B^*|A\rangle\rangle \leq \mathbf{0},$$

which implies that $\langle\langle B^*|A\rangle\rangle = \Phi^*(BA) - \Phi^*(B)\Phi^*(A) = \mathbf{0}$. \square

Lemma A.1 also has the following useful consequence:

Corollary A.2. Let $\Phi^* : \mathcal{L}(\mathcal{K}) \rightarrow \mathcal{L}(\mathcal{H})$ be an operation. Given the sesquilinear mapping defined in Eq. (22), for all $A, B \in \mathcal{L}(\mathcal{K})$ it holds that

$$\|[\Phi^*(A), \Phi^*(B)] - \Phi^*([A, B])\| \leq \|\langle\langle A|A\rangle\rangle\|^{\frac{1}{2}}\|\langle\langle B^*|B^*\rangle\rangle\|^{\frac{1}{2}} + \|\langle\langle A^*|A^*\rangle\rangle\|^{\frac{1}{2}}\|\langle\langle B|B\rangle\rangle\|^{\frac{1}{2}}. \quad (24)$$

Proof. Let us first write

$$[\Phi^*(A), \Phi^*(B)] - \Phi^*([A, B]) = \langle\langle B^*|A\rangle\rangle - \langle\langle A^*|B\rangle\rangle,$$

which gives

$$\|\langle \Phi^*(A), \Phi^*(B) \rangle - \Phi^*([A, B])\| \leq \|\langle B^*|A \rangle\| + \|\langle A^*|B \rangle\|. \quad (25)$$

By Lemma A.1 and the C* identity $\|A^*\| = \|A\| = \|A^*A\|^{\frac{1}{2}} = \|AA^*\|^{\frac{1}{2}}$ for all $A \in \mathcal{L}(\mathcal{H})$, we therefore have

$$\begin{aligned} \|\langle B^*|A \rangle\| &= \|\langle B^*|A \rangle\langle A|B^* \rangle\|^{\frac{1}{2}} \leq \|\langle A|A \rangle\|^{\frac{1}{2}} \|\langle B^*|B^* \rangle\|^{\frac{1}{2}}, \\ \|\langle A^*|B \rangle\| &= \|\langle B|A^* \rangle\langle A^*|B \rangle\|^{\frac{1}{2}} \leq \|\langle A^*|A^* \rangle\|^{\frac{1}{2}} \|\langle B|B \rangle\|^{\frac{1}{2}}. \end{aligned}$$

Inserting the above inequalities in Eq. (25) gives the bound in Eq. (24). \square

Finally, we present the following useful properties of operations:

Lemma A.2. *Let $\Phi^* : \mathcal{L}(\mathcal{K}) \rightarrow \mathcal{L}(\mathcal{H})$ be an operation. For any effects $A \in \mathcal{L}_p(\mathcal{K})$ and $B \in \mathcal{L}_p(\mathcal{H})$, it holds that*

$$\|\Phi^*(A^2) - \Phi^*(A)^2\| \leq 2\|\Phi^*(A) - B\| + \|B - B^2\|.$$

Proof. This inequality (for channels) was given as Eq.(4) in Ref. [88]; the proof below follows Theorem 2 of Ref. [89]. Let us first define $C := \Phi^*(A) - B$ for notational simplicity. Now, given that $\mathbb{0} \leq A \leq \mathbb{1}_{\mathcal{K}}$ implies $A^2 \leq A$, we may write

$$\begin{aligned} \Phi^*(A^2) - \Phi^*(A)^2 &\leq \Phi^*(A) - \Phi^*(A)^2 \\ &= [C, B] + C(\mathbb{1}_{\mathcal{K}} - \Phi^*(A) - B) + B - B^2, \end{aligned}$$

and so we have the bound

$$\begin{aligned} \|\Phi^*(A^2) - \Phi^*(A)^2\| &\leq \|[C, B]\| + \|C(\mathbb{1}_{\mathcal{K}} - \Phi^*(A) - B)\| + \|B - B^2\| \\ &\leq \|[C, B]\| + \|C\| \|\mathbb{1}_{\mathcal{K}} - \Phi^*(A) - B\| + \|B - B^2\| \\ &\leq \|[C, B]\| + \|C\| + \|B - B^2\| \\ &\leq 2\|C\| + \|B - B^2\|. \end{aligned}$$

In the third line we use the fact that A and B are effects which, given that Φ^* is an operation, gives $\mathbb{0} \leq \Phi^*(A) + B \leq 2\mathbb{1}_{\mathcal{H}}$. This in turn implies that $\|\mathbb{1}_{\mathcal{H}} - \Phi^*(A) - B\| \leq 1$. The inequality in the final line follows from Robertson's uncertainty relation, by which we have

$$\begin{aligned} \|[C, B]\| &= \sup_{\|\phi\|=1} |\langle \phi | i[C, B] \phi \rangle| \leq 2\sqrt{\langle \phi | C^2 \phi \rangle - \langle \phi | C \phi \rangle^2} \sqrt{\langle \phi | B^2 \phi \rangle - \langle \phi | B \phi \rangle^2} \\ &\leq 2\|C\| \sqrt{\langle \phi | B^2 \phi \rangle - \langle \phi | B \phi \rangle^2} \\ &\leq \|C\|. \end{aligned}$$

The final line follows from the fact that $\mathbb{0} \leq B \leq \mathbb{1}_{\mathcal{H}}$ implies $\sqrt{\langle \phi | B^2 \phi \rangle - \langle \phi | B \phi \rangle^2} \leq 1/2$. \square

Lemma A.3. *Let $\Phi^* : \mathcal{L}(\mathcal{K}) \rightarrow \mathcal{L}(\mathcal{H})$ be an operation. Assume that $\Phi^*(A) = \mathbb{0}$ for some $A \in \mathcal{L}_p(\mathcal{K})$. It holds that*

$$\Phi^*(AB) = \Phi^*(BA) = \mathbb{0}$$

for all $B \in \mathcal{L}(\mathcal{K})$.

Proof. First, let us note that for any $B \in \mathcal{L}_p(\mathcal{K})$, we have

$$\mathbb{0} \leq \Phi^*(ABA) \leq \|B\| \Phi^*(\sqrt{AA}\sqrt{A}) \leq \|B\| \|A\| \Phi^*(A) = \mathbb{0},$$

and so $\Phi^*(ABA) = \mathbb{0}$. By the two-positivity of CP maps, it follows that for any $B \in \mathcal{L}(\mathcal{K})$ we have

$$\mathbb{0} = \Phi^*(AB^*BA) \geq \Phi^*(AB^*)\Phi^*(BA) \geq \mathbb{0}.$$

The claim immediately follows. \square

B Faithful fixed states and von Neuman algebras

Recall that for channels $\Phi : \mathcal{T}(\mathcal{H}) \rightarrow \mathcal{T}(\mathcal{H})$, and their duals $\Phi^* : \mathcal{L}(\mathcal{H}) \rightarrow \mathcal{L}(\mathcal{H})$, the fixed-point sets are defined as

$$\mathcal{F}(\Phi) := \{T \in \mathcal{T}(\mathcal{H}) : \Phi(T) = T\}, \quad \mathcal{F}(\Phi^*) := \{A \in \mathcal{L}(\mathcal{H}) : \Phi^*(A) = A\}.$$

Lemma B.1 (Lindblad). *Assume that $\mathcal{F}(\Phi)$ contains a faithful state. Then $\mathcal{F}(\Phi^*)$ is a von Neumann algebra.*

Proof. Suppose $B \in \mathcal{F}(\Phi^*)$, and define the operator $\Phi^*(B^*B) - \Phi^*(B^*)\Phi^*(B) = \Phi^*(B^*B) - B^*B$, which is positive due to the two-positivity of CP maps. Let $\mathcal{F}(\Phi)$ contain a faithful state ω . Then we have

$$\text{tr}[\omega(\Phi^*(B^*B) - B^*B)] = \text{tr}[\omega(B^*B - B^*B)] = 0.$$

But since ω is faithful and $\Phi^*(B^*B) - B^*B$ is positive, $\text{tr}[\omega(\Phi^*(B^*B) - B^*B)] = 0$ implies that $\Phi^*(B^*B) = B^*B$. Corollary A.1 therefore implies that for all $A \in \mathcal{L}(\mathcal{H})$,

$$\Phi^*(AB) = \Phi^*(A)B.$$

Therefore, if $A \in \mathcal{F}(\Phi^*)$, then $\Phi^*(AB) = AB$, and so $\mathcal{F}(\Phi^*)$ is closed under multiplication and is therefore a $*$ -algebra. Finally, if $\mathcal{F}(\Phi^*)$ is an algebra, then $\mathcal{F}(\Phi^*) = \{K_i, K_i^*\}' := \{A \in \mathcal{L}(\mathcal{H}) : [K_i, A] = [K_i^*, A] = \mathbf{0} \forall i\}$, with $\{K_i\}$ any Kraus representation of Φ [90], making $\mathcal{F}(\Phi^*)$ a von Neumann algebra (as the commutant of a self-adjoint subset of $\mathcal{L}(\mathcal{H})$) [55]. \square

If $\mathcal{F}(\Phi^*)$ is a von Neumann algebra, it holds that for any self-adjoint operator $A \in \mathcal{F}(\Phi^*)$, the spectral measure of A is also contained in $\mathcal{F}(\Phi^*)$. In the case that A has a discrete spectrum, i.e., $A = \sum_n \lambda_n P_n$, this implies that $\{P_n\} \subset \mathcal{F}(\Phi^*)$.

C Fixed points of instrument channels

Here we prove a useful result regarding the fixed-point structure of the E-channel $\mathcal{I}_{\mathcal{X}}^*$, describing a non-selective measurement of an observable E , which we shall use in several places in this paper.

Lemma C.1. *Let \mathcal{I} be an instrument compatible with an observable E acting in \mathcal{H}_S . The following hold:*

- (i) *If E is sharp, then $\mathcal{F}(\mathcal{I}_{\mathcal{X}}^*) \subset \mathcal{I}_{\mathcal{X}}^*(\mathcal{L}(\mathcal{H}_S)) \subset E'$.*
- (ii) *If $\mathcal{F}(\mathcal{I}_{\mathcal{X}}^*)$ is a von Neumann algebra, then $\mathcal{F}(\mathcal{I}_{\mathcal{X}}^*) \subset E'$.*
- (iii) *If $\mathcal{I}_{\mathcal{X}}$ fully conserves a self-adjoint operator $A \in \mathcal{L}_s(\mathcal{H}_S)$, then $A \in E'$.*

Proof. All E-compatible instruments \mathcal{I} admit a measurement scheme $\mathcal{M} := (\mathcal{H}_A, \xi, \mathcal{E}, Z)$. Therefore, by the channel $\Gamma_{\xi}^{\mathcal{E}}$ defined in Eq. (3), we may write $[E(x), \mathcal{I}_{\mathcal{X}}^*(A)] = [\Gamma_{\xi}^{\mathcal{E}}(\mathbf{1}_S \otimes Z(x)), \Gamma_{\xi}^{\mathcal{E}}(A \otimes \mathbf{1}_A)]$. Since $[\mathbf{1}_S \otimes Z(x), A \otimes \mathbf{1}_A] = \mathbf{0}$, and $Z(x)$ are positive operators, then by the sesquilinear mapping $\langle\langle A|B \rangle\rangle := \Gamma_{\xi}^{\mathcal{E}}(A^*B) - \Gamma_{\xi}^{\mathcal{E}}(A^*)\Gamma_{\xi}^{\mathcal{E}}(B)$ and Corollary A.2 we obtain

$$\|[E(x), \mathcal{I}_{\mathcal{X}}^*(A)]\| \leq \|\langle\langle \mathbf{1}_S \otimes Z(x) | \mathbf{1}_S \otimes Z(x) \rangle\rangle\|^{\frac{1}{2}} \left(\|\langle\langle A \otimes \mathbf{1}_A | A \otimes \mathbf{1}_A \rangle\rangle\|^{\frac{1}{2}} + \|\langle\langle A^* \otimes \mathbf{1}_A | A^* \otimes \mathbf{1}_A \rangle\rangle\|^{\frac{1}{2}} \right). \quad (26)$$

Since $Z(x)$ are effects, it follows that

$$\|\langle\langle \mathbf{1}_S \otimes Z(x) | \mathbf{1}_S \otimes Z(x) \rangle\rangle\| = \Gamma_{\xi}^{\mathcal{E}}(\mathbf{1}_S \otimes Z(x)^2) - \Gamma_{\xi}^{\mathcal{E}}(\mathbf{1}_S \otimes Z(x))^2 \leq \Gamma_{\xi}^{\mathcal{E}}(\mathbf{1}_S \otimes Z(x)) - \Gamma_{\xi}^{\mathcal{E}}(\mathbf{1}_S \otimes Z(x))^2 = E(x) - E(x)^2.$$

On the other hand, we have $\langle\langle A \otimes \mathbf{1}_A | A \otimes \mathbf{1}_A \rangle\rangle = \mathcal{I}_{\mathcal{X}}^*(A^*A) - \mathcal{I}_{\mathcal{X}}^*(A^*)\mathcal{I}_{\mathcal{X}}^*(A)$ and $\langle\langle A^* \otimes \mathbf{1}_A | A^* \otimes \mathbf{1}_A \rangle\rangle = \mathcal{I}_{\mathcal{X}}^*(AA^*) - \mathcal{I}_{\mathcal{X}}^*(A)\mathcal{I}_{\mathcal{X}}^*(A^*)$. We thus obtain from Eq. (26) the bound

$$\|[E(x), \mathcal{I}_{\mathcal{X}}^*(A)]\| \leq \|E(x) - E(x)^2\|^{\frac{1}{2}} \left(\|\mathcal{I}_{\mathcal{X}}^*(A^*A) - \mathcal{I}_{\mathcal{X}}^*(A^*)\mathcal{I}_{\mathcal{X}}^*(A)\|^{\frac{1}{2}} + \|\mathcal{I}_{\mathcal{X}}^*(AA^*) - \mathcal{I}_{\mathcal{X}}^*(A)\mathcal{I}_{\mathcal{X}}^*(A^*)\|^{\frac{1}{2}} \right). \quad (27)$$

Now we may prove (i). If E is sharp, then the upper bound of Eq. (27) vanishes and so for all $A \in \mathcal{L}(\mathcal{H}_S)$, $\mathcal{I}_{\mathcal{X}}^*(A) \in E'$. As such, $\mathcal{I}_{\mathcal{X}}^*(\mathcal{L}(\mathcal{H}_S)) \subset E'$. That $\mathcal{F}(\mathcal{I}_{\mathcal{X}}^*) \subset \mathcal{I}_{\mathcal{X}}^*(\mathcal{L}(\mathcal{H}_S))$ is trivial.

Now we prove (ii). Assume that $A \in \mathcal{F}(\mathcal{I}_{\mathcal{X}}^*)$, which implies that $A^* \in \mathcal{F}(\mathcal{I}_{\mathcal{X}}^*)$. But if $\mathcal{F}(\mathcal{I}_{\mathcal{X}}^*)$ is a von Neumann algebra, this implies that $A^*A, AA^* \in \mathcal{F}(\mathcal{I}_{\mathcal{X}}^*)$, and so the upper bound of Eq. (27) vanishes. Consequently, we see that for all $A \in \mathcal{L}(\mathcal{H}_S)$, $A \in \mathcal{F}(\mathcal{I}_{\mathcal{X}}^*) \implies A \in E'$, which implies that $\mathcal{F}(\mathcal{I}_{\mathcal{X}}^*) \subset E'$.

Finally, let us prove (iii). Let A be a self-adjoint operator, and assume that $\mathcal{I}_{\mathcal{X}}$ fully conserves A . By Definition 2 it holds that $\mathcal{I}_{\mathcal{X}}^*(A^k) = A^k$ for $k = 1, 2$, and so once again the upper bound of Eq. (27) vanishes, implying that $A \in E'$. \square

D Disturbance, commutation, and compatibility

The pair of observables $E := \{E(x) : x \in \mathcal{X}\}$ and $F := \{F(y) : y \in \mathcal{Y}\}$ acting in \mathcal{H}_s are *compatible*, or *jointly measurable*, if they admit a joint observable $G := \{G(x, y) : (x, y) \in \mathcal{X} \times \mathcal{Y}\}$ so that

$$\sum_{y \in \mathcal{Y}} G(x, y) = E(x), \quad \sum_{x \in \mathcal{X}} G(x, y) = F(y) \quad \forall x \in \mathcal{X}, y \in \mathcal{Y}. \quad (28)$$

If E and F do not admit a joint observable, then they are *incompatible* [21]. Now let \mathcal{I} be an E -compatible instrument, and assume that $F \subset \mathcal{F}(\mathcal{I}_\mathcal{X}^*)$. In such a case, we may choose G as $G(x, y) = \mathcal{I}_x^*(F(y))$, which satisfies Eq. (28). It follows that non-disturbance implies compatibility, and so for two incompatible observables E and F , no E -instrument \mathcal{I} exists that satisfies $F \subset \mathcal{F}(\mathcal{I}_\mathcal{X}^*)$. Note that while non-disturbance requires compatibility, compatibility does not guarantee non-disturbance. For instance, while any observable is compatible with itself, for every informationally complete observable the fixed-point set of its compatible channel is trivial. Indeed, the size of the fixed-point set of an E -channel is strongly related to the amount of information given by E as shown in Ref. [16]. Furthermore, as shown in Ref. [4], there exist pairs of compatible observables E and F where E admits an instrument that does not disturb F , but all possible F -instruments necessarily disturb E . This further demonstrates that unlike compatibility, non-disturbance is not symmetric.

As shown in Ref. [88], the pair of observables E and F are compatible only if

$$\| [E(x), F(y)] \| \leq 2 \| E(x) - E(x)^2 \|^{1/2} \| F(y) - F(y)^2 \|^{1/2} \quad \forall x \in \mathcal{X}, y \in \mathcal{Y}. \quad (29)$$

Commutation is a sufficient condition for compatibility; if E commutes with F , then there is a joint observable G with effects $G(x, y) = E(x)F(y) \equiv (\sqrt{E(x)}\sqrt{F(y)})^*(\sqrt{E(x)}\sqrt{F(y)})$. On the other hand, if either E or F is sharp, in which case the upper bound of Eq. (29) vanishes, then commutation is a necessary condition for compatibility [91]. For two non-commuting observables to be compatible, therefore, their effects must be sufficiently unsharp. We now provide a bound for the disturbance of F by an E -instrument \mathcal{I} , in terms of the commutation between the effects of E and F .

Proposition D.1. *Consider the observables E and F acting in \mathcal{H}_s , and let $\|\delta(y)\|$ be the disturbance of the effects of F caused by an E -instrument \mathcal{I} . Then for all $x \in \mathcal{X}$ and $y \in \mathcal{Y}$ it holds that*

$$\| [E(x), F(y)] \| \leq \|\delta(y)\| + 2 \| E(x) - E(x)^2 \|^{1/2} \| \mathcal{I}_\mathcal{X}^*(F(y)^2) - \mathcal{I}_\mathcal{X}^*(F(y))^2 \|^{1/2}. \quad (30)$$

If F is non-disturbed by \mathcal{I} , that is, if $\delta = 0$, then for all $x \in \mathcal{X}$ and $y \in \mathcal{Y}$ it holds that

$$\| [E(x), F(y)] \| \leq 2 \| E(x) - E(x)^2 \|^{1/2} \| \mathcal{I}_\mathcal{X}^*(F(y)^2) - F(y)^2 \|^{1/2}. \quad (31)$$

Proof. By Eq. (8), we may write

$$[E(x), F(y)] = [\delta(y), E(x)] + [E(x), \mathcal{I}_\mathcal{X}^*(F(y))]. \quad (32)$$

Every E -instrument \mathcal{I} admits a measurement scheme $\mathcal{M} := (\mathcal{H}_\mathcal{A}, \xi, \mathcal{E}, \mathcal{Z})$. Using the channel $\Gamma_\xi^\mathcal{E}$ defined in Eq. (3), we may therefore write $[E(x), \mathcal{I}_\mathcal{X}^*(F(y))] = [\Gamma_\xi^\mathcal{E}(\mathbb{1}_s \otimes Z(x)), \Gamma_\xi^\mathcal{E}(F(y) \otimes \mathbb{1}_\mathcal{A})]$. Given that $[\mathbb{1}_s \otimes Z(x), F(y) \otimes \mathbb{1}_\mathcal{A}] = \mathbb{O}$, then by the sesquilinear mapping $\langle\langle A|B \rangle\rangle := \Gamma_\xi^\mathcal{E}(A^*B) - \Gamma_\xi^\mathcal{E}(A^*)\Gamma_\xi^\mathcal{E}(B)$ and Corollary A.2, we obtain from Eq. (32) the bound

$$\| [E(x), F(y)] \| \leq \| [\delta(y), E(x)] \| + 2 \| \langle\langle \mathbb{1}_s \otimes Z(x) | \mathbb{1}_s \otimes Z(x) \rangle\rangle \|^{1/2} \| \langle\langle F(y) \otimes \mathbb{1}_\mathcal{A} | F(y) \otimes \mathbb{1}_\mathcal{A} \rangle\rangle \|^{1/2}. \quad (33)$$

Since $E(x)$ is an effect, then as shown in Lemma A.2 we have $\| [\delta(y), E(x)] \| \leq \|\delta(y)\|$. As shown in Lemma C.1, we have $\langle\langle \mathbb{1}_s \otimes Z(x) | \mathbb{1}_s \otimes Z(x) \rangle\rangle \leq E(x) - E(x)^2$ and $\langle\langle F(y) \otimes \mathbb{1}_\mathcal{A} | F(y) \otimes \mathbb{1}_\mathcal{A} \rangle\rangle = \mathcal{I}_\mathcal{X}^*(F(y)^2) - \mathcal{I}_\mathcal{X}^*(F(y))^2$. We therefore obtain from Eq. (33) the bound given in Eq. (30). If F is non-disturbed by \mathcal{I} , then $\|\delta(y)\| = 0$ and $\mathcal{I}_\mathcal{X}^*(F(y)^2) = F(y)^2$ for all y . We thus arrive at Eq. (31). \square

We see that when E commutes with F the lower bound of Eq. (31) vanishes, in which case Proposition D.1 does not prohibit non-disturbance. Indeed, in the case of commuting observables there always exists a non-disturbing instrument; since $E' \subset \mathcal{F}(\mathcal{I}_\mathcal{X}^L)$ always holds, where \mathcal{I}^L is the Lüders E -instrument defined in Eq. (1), then a Lüders measurement of E is guaranteed not to disturb any F commuting with E [63]. On the other hand, if E does not commute with F , then Proposition D.1 allows us to obtain a lower bound for the disturbance that results given any E -compatible instrument, determined only by the unsharpness and non-commutation of E and F :

Corollary D.1. *Consider the setup of Proposition D.1. For all $x \in \mathcal{X}$ and $y \in \mathcal{Y}$, it also holds that*

$$\|[\mathbf{E}(x), \mathbf{F}(y)]\| \leq \|\delta(y)\| + 2\|\mathbf{E}(x) - \mathbf{E}(x)^2\|^{\frac{1}{2}} \left(2\|\delta(y)\| + \|\mathbf{F}(y) - \mathbf{F}(y)^2\| \right)^{\frac{1}{2}}. \quad (34)$$

Proof. Since $\mathbf{F}(y)$ are effects and $\mathcal{I}_{\mathcal{X}}^*$ is a channel, then by Lemma A.2 we have $\|\mathcal{I}_{\mathcal{X}}^*(\mathbf{F}(y)^2) - \mathcal{I}_{\mathcal{X}}^*(\mathbf{F}(y))^2\| \leq 2\|\delta(y)\| + \|\mathbf{F}(y) - \mathbf{F}(y)^2\|$. As such, Eq. (34) is obtained directly from Eq. (30). \square

Note that while Corollary D.1 provides a lower bound for the disturbance, which is strictly positive whenever either \mathbf{E} or \mathbf{F} is sharp and these observables do not commute, such a lower bound will differ depending on whether \mathbf{E} or \mathbf{F} is sharp; if \mathbf{E} is sharp, we have $\delta \geq \max_{x,y} \|[\mathbf{E}(x), \mathbf{F}(y)]\|$, whereas if \mathbf{F} is sharp but \mathbf{E} is unsharp, the lower bound for the disturbance may be smaller. Let us illustrate this with the following example. Consider a system $\mathcal{H}_s \simeq \mathbb{C}^2$, with the orthonormal basis $\{|0\rangle, |1\rangle\}$, and define $|\pm\rangle := \frac{1}{\sqrt{2}}(|0\rangle \pm |1\rangle)$. Now consider a pair of binary observables $\mathbf{A} = \{\mathbf{A}(a) : a = 0, 1\}$ and $\mathbf{B}_\lambda = \{\mathbf{B}_\lambda(b) : b = \pm\}$ acting in \mathcal{H}_s , defined by $\mathbf{A}(a) = |a\rangle\langle a|$ and $\mathbf{B}_\lambda(b) = \lambda|b\rangle\langle b| + (1 - \lambda)\frac{\mathbb{1}}{2}$ for some $0 \leq \lambda \leq 1$. It is simple to verify that $\|[\mathbf{A}(a), \mathbf{B}_\lambda(b)]\| = \frac{\lambda}{2}$ for any $a = 0, 1$ and $b = \pm$. Now we may evaluate the disturbance of one of these observables caused by a Lüders measurement of the other. The disturbance of $\mathbf{B}_\lambda(b)$ by a Lüders measurement of \mathbf{A} reads $\|\delta(b)\| = \frac{\lambda}{2}$ for each b . Since \mathbf{A} is sharp, then by setting $\mathbf{E} = \mathbf{A}$ and $\mathbf{F} = \mathbf{B}_\lambda$, we see that the inequality in Eq. (34) is tight. On the other hand, the disturbance of $\mathbf{A}(a)$ by a Lüders measurement of \mathbf{B}_λ reads $\|\delta(a)\| = \frac{1 - \sqrt{1 - \lambda^2}}{2}$ for each a , which is smaller than $\frac{\lambda}{2}$ for $0 < \lambda < 1$.

Let us now consider the case of non-disturbance more carefully. First, let us note that when we set $\|\delta(y)\| = 0$, Eq. (34) reduces to the compatibility bound of Eq. (29), and states that for non-disturbance to be possible when \mathbf{E} and \mathbf{F} do not commute, then both observables must be sufficiently unsharp so as to be compatible. To be sure, compatibility is a necessary condition for non-disturbance, and the fact that Eq. (34) does not contradict the compatibility bound is not surprising. On the other hand, in the case of non-disturbance this bound is also not very informative—it is possible for two observables to be compatible, while a measurement of one still disturbs the other. To gain a better understanding of non-disturbance, let us consider instead Eq. (31), the upper bound of which is smaller than the upper bound in Eq. (34) when we set $\|\delta(y)\| = 0$, and vanishes if both $\mathbf{F} \subset \mathcal{F}(\mathcal{I}_{\mathcal{X}}^*)$ and $\mathbf{F}^2 := \{\mathbf{F}(y)^2 : y \in \mathcal{Y}\} \subset \mathcal{F}(\mathcal{I}_{\mathcal{X}}^*)$ hold. We immediately see that while unsharpness of both \mathbf{E} and \mathbf{F} is necessary for non-disturbance when \mathbf{E} and \mathbf{F} do not commute, it is not sufficient; as shown in Ref. [4] there are at least two classes of unsharp observables \mathbf{F} where given *any* instrument \mathcal{I} , i.e., including instruments that measure an unsharp observable \mathbf{E} that does not commute with \mathbf{F} but is still compatible with \mathbf{F} , it holds that $\mathbf{F} \subset \mathcal{F}(\mathcal{I}_{\mathcal{X}}^*)$ guarantees $\mathbf{F}^2 \subset \mathcal{F}(\mathcal{I}_{\mathcal{X}}^*)$: if \mathbf{F} is a rank-1 observable, or if \mathbf{F} is an “informationally equivalent coarse-graining” of a sharp observable.

Let us consider the first option. If \mathbf{F} is a rank-1 observable, then all the effects of \mathbf{F} may be written as $\mathbf{F}(y) = \lambda_y P_y$, where P_y is a rank-1 projection operator and $\lambda_y \in (0, 1]$. As shown in [92], all observables \mathbf{E} that are compatible with a rank-1 observable \mathbf{F} are the post-processings of \mathbf{F} , that is, the effects of \mathbf{E} may be written as $\mathbf{E}(x) = \sum_y p(x|y) \mathbf{F}(y)$, where $\{p(x|y)\}$ is a family of non-negative numbers satisfying $\sum_x p(x|y) = 1$ for all y . It follows that so long as \mathbf{F} is a non-commutative rank-1 observable, then there exists an unsharp observable \mathbf{E} that is compatible with \mathbf{F} but does not commute with \mathbf{F} . But note that $\mathcal{I}_{\mathcal{X}}^*(\mathbf{F}(y)) = \mathbf{F}(y)$ if and only if $\mathcal{I}_{\mathcal{X}}^*(P_y) = P_y$. As such, $\mathcal{I}_{\mathcal{X}}^*(\mathbf{F}(y)^2) = \lambda_y^2 \mathcal{I}_{\mathcal{X}}^*(P_y) = \lambda_y^2 P_y = \mathbf{F}(y)^2$. It follows that \mathbf{F} will be non-disturbed by an \mathbf{E} -compatible instrument \mathcal{I} only if \mathbf{E} commutes with \mathbf{F} .

Let us now consider the second option. We say that \mathbf{F} is an informationally equivalent coarse-graining of a sharp observable $\mathbf{G} := \{\mathbf{G}(z) : z \in \mathcal{Z}\}$ if there exists an invertible stochastic matrix M such that

$$\mathbf{F}(y) = \sum_z M_{y,z} \mathbf{G}(z), \quad \mathbf{G}(z) = \sum_y M_{z,y}^{-1} \mathbf{F}(y).$$

\mathbf{F} and \mathbf{G} are informationally equivalent because a measurement of \mathbf{F} produces different probability distributions for two states ρ_1 and ρ_2 if and only if these states produce different probability distributions given a measurement of \mathbf{G} . Since \mathbf{G} is sharp, then $\mathbf{F}(y)^2 = \sum_z M_{y,z}^2 \mathbf{G}(z)$. Now assume that $\mathbf{F} \subset \mathcal{F}(\mathcal{I}_{\mathcal{X}}^*)$. It is simple to verify that this implies $\mathbf{G} \subset \mathcal{F}(\mathcal{I}_{\mathcal{X}}^*)$. Therefore, we have $\mathcal{I}_{\mathcal{X}}^*(\mathbf{F}(y)^2) = \sum_z M_{y,z}^2 \mathcal{I}_{\mathcal{X}}^*(\mathbf{G}(z)) = \sum_z M_{y,z}^2 \mathbf{G}(z) = \mathbf{F}(y)^2$. Once again, \mathbf{F} will be non-disturbed by an \mathbf{E} -compatible instrument \mathcal{I} only if \mathbf{E} commutes with \mathbf{F} .

Both of the above examples offer a very simple interpretation in terms of compatibility. If \mathbf{F} is a rank-1 observable, then non-disturbance of \mathbf{F} implies non-disturbance of sharp rank-1 effects P_y . Since non-disturbance requires compatibility, this implies that \mathbf{E} must commute with all P_y , and hence with \mathbf{F} . On the other hand, if \mathbf{F} is a classical coarse-graining of a sharp observable \mathbf{G} , then non-disturbance of \mathbf{F} implies non-disturbance of \mathbf{G} , and by compatibility \mathbf{E} must commute with \mathbf{G} . Since the effects of \mathbf{F} are constructed as a mixture of the (projective) effects of \mathbf{G} , this concludes that \mathbf{E} must commute with \mathbf{F} .

E Properties of repeatable instruments

In this section, we prove a series of useful results regarding the structure of repeatable instruments, and the measurement schemes that implement them.

Proposition E.1. *Let $\mathcal{M} := (\mathcal{H}_A, \xi, \mathcal{E}, \mathcal{Z})$ be a measurement scheme for an \mathbf{E} -compatible instrument \mathcal{I} acting in \mathcal{H}_S . If \mathcal{I} is repeatable, then the following hold:*

- (i) *For all $x \in \mathcal{X}$ and $n \in \mathbb{N}$, it holds that $\mathbf{E}(x) = \Gamma_\xi^\mathcal{E}(\mathbf{E}(x)^n \otimes \mathbf{1}_A) = \Gamma_\xi^\mathcal{E}(\mathbf{1}_S \otimes \mathbf{Z}(x)^n)$.*
- (ii) *For all $x \in \mathcal{X}$, it holds that $\mathbf{E}(x)$ and $\mathbf{Z}(x)$ have 1 as an eigenvalue, and so there exist projection operators $\mathbf{P}(x) \in \mathcal{L}_p(\mathcal{H}_S)$ and $\mathbf{Q}(x) \in \mathcal{L}_p(\mathcal{H}_A)$ which project onto the eigenvalue-1 eigenspaces of $\mathbf{E}(x)$ and $\mathbf{Z}(x)$, respectively.*
- (iii) *For all $x, y \in \mathcal{X}$, it holds that $\mathbf{P}(x)\mathbf{E}(y) = \mathbf{P}(x)\mathbf{P}(y) = \delta_{x,y}\mathbf{P}(x)$ and $\mathbf{Q}(x)\mathbf{Z}(y) = \mathbf{Q}(x)\mathbf{Q}(y) = \delta_{x,y}\mathbf{Q}(x)$.*
- (iv) *For all $x \in \mathcal{X}$ and $A \in \mathcal{L}(\mathcal{H}_S \otimes \mathcal{H}_A)$, it holds that $\Gamma_\xi^\mathcal{E}(\mathbf{P}^c(x) \otimes \mathbf{1}_A) = \Gamma_\xi^\mathcal{E}(A\mathbf{P}^c(x) \otimes \mathbf{1}_A) = \mathbf{0}$ and $\Gamma_\xi^\mathcal{E}(\mathbf{1}_S \otimes \mathbf{Q}^c(x)A) = \Gamma_\xi^\mathcal{E}(A\mathbf{1}_S \otimes \mathbf{Q}^c(x)) = \mathbf{0}$, where $\mathbf{P}^c(x) := \mathbf{E}(x) - \mathbf{P}(x)$ and $\mathbf{Q}^c(x) := \mathbf{Z}(x) - \mathbf{Q}(x)$.*
- (v) *For all $A \in \mathcal{L}(\mathcal{H}_S \otimes \mathcal{H}_A)$, it holds that $\Gamma_\xi^\mathcal{E}(A) = \Gamma_\xi^\mathcal{E}(\mathbf{R}A\mathbf{R})$, where $\mathbf{R} := \sum_{x \in \mathcal{X}} \mathbf{P}(x) \otimes \mathbf{Q}(x)$.*
- (vi) *For all $x \in \mathcal{X}$, $A \in \mathcal{L}(\mathcal{H}_S)$, and $B \in \mathcal{L}(\mathcal{H}_A)$, it holds that $\mathcal{I}_\mathcal{X}^*(\mathbf{E}(x)A) = \mathcal{I}_\mathcal{X}^*(A\mathbf{E}(x)) = \mathcal{I}_\mathcal{X}^*(\mathbf{P}(x)A\mathbf{P}(x))$ and $\Lambda^*(\mathbf{Z}(x)B) = \Lambda^*(B\mathbf{Z}(x)) = \Lambda^*(\mathbf{Q}(x)B\mathbf{Q}(x))$.*
- (vii) *For all $x \in \mathcal{X}$ and $A \in \mathcal{L}(\mathcal{H}_S)$, it holds that $\mathcal{I}_x^*(A) = \mathcal{I}_x^*(\mathbf{P}(x)A\mathbf{P}(x))$.*

Proof. (i): The repeatability condition implies that for all $x \in \mathcal{X}$, it holds that $\mathbf{E}(x) = \mathcal{I}_x^*(\mathbf{E}(x)) = \Gamma_\xi^\mathcal{E}(\mathbf{E}(x) \otimes \mathbf{Z}(x))$. It follows that for any state $\rho \in \mathcal{S}(\mathcal{H}_S)$, we have

$$\begin{aligned} \text{tr}[\rho\mathbf{E}(x)] &= \text{tr}[\rho\Gamma_\xi^\mathcal{E}(\mathbf{E}(x) \otimes \mathbf{Z}(x))] \\ &\leq \text{tr}[\rho\Gamma_\xi^\mathcal{E}(\mathbf{E}(x)^2 \otimes \mathbf{1}_A)]^{\frac{1}{2}} \text{tr}[\rho\Gamma_\xi^\mathcal{E}(\mathbf{1}_S \otimes \mathbf{Z}(x)^2)]^{\frac{1}{2}} \\ &\leq \text{tr}[\rho\Gamma_\xi^\mathcal{E}(\mathbf{E}(x) \otimes \mathbf{1}_A)]^{\frac{1}{2}} \text{tr}[\rho\Gamma_\xi^\mathcal{E}(\mathbf{1}_S \otimes \mathbf{Z}(x))]^{\frac{1}{2}} \\ &= \text{tr}[\rho\mathbf{E}(x)]. \end{aligned}$$

Here, the second line follows from the Cauchy-Schwarz inequality, the third line follows from the fact that $\mathbf{E}(x)$ and $\mathbf{Z}(x)$ are effects and so $\mathbf{E}(x)^2 \leq \mathbf{E}(x)$ and $\mathbf{Z}(x)^2 \leq \mathbf{Z}(x)$, and the final line follows from the fact that repeatability implies first-kindness and that \mathcal{M} is a measurement scheme for \mathbf{E} . As the second inequality must be an equality, we thus have $\mathbf{E}(x) = \Gamma_\xi^\mathcal{E}(\mathbf{E}(x)^n \otimes \mathbf{1}_A) = \Gamma_\xi^\mathcal{E}(\mathbf{1}_S \otimes \mathbf{Z}(x)^n)$ for $n = 1, 2$. To show that the relations hold for all $n \in \mathbb{N}$, it suffices to show that for all ρ , the Cauchy-Schwarz inequality and the above arguments implies

$$\begin{aligned} 0 \leq \text{tr}[\rho\Gamma_\xi^\mathcal{E}((\mathbf{E}(x)^n - \mathbf{E}(x)^{n+1}) \otimes \mathbf{1}_A)] &\leq \text{tr}[\rho\Gamma_\xi^\mathcal{E}(\mathbf{E}(x)^{2(n-1)} \otimes \mathbf{1}_A)]^{\frac{1}{2}} \text{tr}[\rho\Gamma_\xi^\mathcal{E}((\mathbf{E}(x) - \mathbf{E}(x)^2) \otimes \mathbf{1}_A)]^{\frac{1}{2}} \\ &\leq \text{tr}[\rho\Gamma_\xi^\mathcal{E}(\mathbf{E}(x)^{2(n-1)} \otimes \mathbf{1}_A)]^{\frac{1}{2}} \text{tr}[\rho\Gamma_\xi^\mathcal{E}((\mathbf{E}(x) - \mathbf{E}(x)^2) \otimes \mathbf{1}_A)]^{\frac{1}{2}} = 0, \end{aligned}$$

and so it holds that $\Gamma_\xi^\mathcal{E}((\mathbf{E}(x)^n - \mathbf{E}(x)^{n+1}) \otimes \mathbf{1}_A) = \mathbf{0}$. Similar steps show that $\Gamma_\xi^\mathcal{E}(\mathbf{1}_S \otimes (\mathbf{Z}(x)^n - \mathbf{Z}(x)^{n+1})) = \mathbf{0}$. The claims are thus obtained by induction.

- (ii): Note that for any operation $\Phi^* : \mathcal{L}(\mathcal{K}) \rightarrow \mathcal{L}(\mathcal{H})$, it holds that $\|\Phi^*(A)\| \leq \|A\|$ for all $A \in \mathcal{L}(\mathcal{K})$. As such, by (i) we have $\|\mathbf{E}(x)\| = \|\Gamma_\xi^\mathcal{E}(\mathbf{E}(x)^2 \otimes \mathbf{1}_A)\| \leq \|\mathbf{E}(x)^2\| = \|\mathbf{E}(x)\|^2$. But since $\mathbf{E}(x)$ is an effect it also holds that $\|\mathbf{E}(x)\| \geq \|\mathbf{E}(x)\|^2$. It follows that $\|\mathbf{E}(x)\|$ is either zero or one. As we assume that $\mathbf{E}(x)$ is not vanishing, then $\|\mathbf{E}(x)\| = 1$ follows. Similarly, we have $1 = \|\mathbf{E}(x)\| = \|\Gamma_\xi^\mathcal{E}(\mathbf{1}_S \otimes \mathbf{Z}(x))\| \leq \|\mathbf{Z}(x)\|$, and since $\mathbf{Z}(x)$ is an effect, then it must hold that $\|\mathbf{Z}(x)\| = 1$.

Now we shall show that $\mathbf{E}(x)$ has 1 as an eigenvalue, i.e., there exists a unit-vector $\psi \in \mathcal{H}_S$ such that $\mathbf{E}(x)\psi = \psi$. If this is not so, then we would have $\lim_{n \rightarrow \infty} \mathbf{E}(x)^n = \mathbf{0}$, which would contradict (i). Therefore, there exists a projection operator $\mathbf{P}(x)$ that projects onto the eigenvalue-1 eigenspace of $\mathbf{E}(x)$. Similar arguments hold for $\mathbf{Z}(x)$ and $\mathbf{Q}(x)$.

- (iii): For each x , define $\mathbf{P}^c(x) := \mathbf{E}(x) - \mathbf{P}(x)$. Since $\mathbf{E}(x)$ is an effect and $\mathbf{P}(x)$ projects onto the eigenvalue-1 eigenspace of $\mathbf{E}(x)$, it trivially holds that $\psi \in \text{supp}(\mathbf{P}(x)) \implies \psi \in \ker(\mathbf{P}^c(x))$. Now, given that

$\psi \in \text{supp}(P(x))$ implies that $P(x)\psi = \psi$, and denoting the null vector in \mathcal{H}_s as \emptyset , we have

$$\begin{aligned}\emptyset &= (\mathbf{1}_s - P(x))\psi \\ &= (\mathbf{1}_s - E(x))\psi + P^c(x)\psi \\ &= (\mathbf{1}_s - E(x))\psi \\ &= \sum_{y \neq x} E(y)\psi.\end{aligned}$$

By positivity of $E(y)$, the above equation implies that

$$\sum_{y \neq x} \langle \psi | E(y)\psi \rangle = \sum_{y \neq x} \langle \sqrt{E(y)}\psi | \sqrt{E(y)}\psi \rangle = 0,$$

which can be satisfied only if $\sqrt{E(y)}\psi = \emptyset \implies E(y)\psi = \emptyset$ for all $y \neq x$. We thus have $\psi \in \text{supp}(P(x)) \implies \psi \in \ker(E(y)) \forall y \neq x$, and so the support of $P(x)$ must be orthogonal to the support of $E(y)$ for all $y \neq x$. That $P(x)$ and $P(y)$ for $x \neq y$ have orthogonal supports follows trivially. Similar arguments hold for $Q(x)$, $Z(y)$, and $Q(y)$.

- (iv): It trivially holds that $\|P^c(x)\| < 1$ and $\|Q^c(x)\| < 1$. We thus have $\lim_{n \rightarrow \infty} P^c(x)^n = \mathbf{0}$ and $\lim_{n \rightarrow \infty} Q^c(x)^n = \mathbf{0}$. As stated in (iii), the supports of $P(x)$ and $P^c(x)$ are orthogonal, and so it holds that $P(x)P^c(x) = P^c(x)P(x) = \mathbf{0}$. As such, for all $n \in \mathbb{N}$ we have $E(x)^n = P(x) + P^c(x)^n$. Similarly, $Z(x)^n = Q(x) + Q^c(x)^n$. By (i), it holds that $\Gamma_\xi^\mathcal{E}(E(x)^n \otimes \mathbf{1}_A) = \Gamma_\xi^\mathcal{E}(P(x) \otimes \mathbf{1}_A) + \Gamma_\xi^\mathcal{E}(P^c(x)^n \otimes \mathbf{1}_A) = E(x)$ for all $n \in \mathbb{N}$, and so it must hold that $\Gamma_\xi^\mathcal{E}(P^c(x) \otimes \mathbf{1}_A) = \lim_{n \rightarrow \infty} \Gamma_\xi^\mathcal{E}(P^c(x)^n \otimes \mathbf{1}_A) = \mathbf{0}$. Similarly, $E(x) = \Gamma_\xi^\mathcal{E}(\mathbf{1}_s \otimes Q(x)) + \Gamma_\xi^\mathcal{E}(\mathbf{1}_s \otimes Q^c(x)^n)$ for all $n \in \mathbb{N}$ implies that $\Gamma_\xi^\mathcal{E}(\mathbf{1}_s \otimes Q^c(x)) = \mathbf{0}$. The claim follows from Lemma A.3.
- (v): First, let us note that by repeatability, it holds that $\Gamma_\xi^\mathcal{E}(E(x) \otimes Z(y)) = \mathcal{I}_y^*(E(x)) = \mathbf{0}$ for all $x \neq y$. By Lemma A.3, it follows that $\Gamma_\xi^\mathcal{E}(E(x) \otimes Z(y)A) = \Gamma_\xi^\mathcal{E}(AE(x) \otimes Z(y)) = \mathbf{0}$ holds for all A and $x \neq y$. We may therefore write

$$\begin{aligned}\Gamma_\xi^\mathcal{E}(A) &= \sum_{x, x', y, y'} \Gamma_\xi^\mathcal{E}(E(x) \otimes Z(x')AE(y) \otimes Z(y')) \\ &= \sum_{x, y} \Gamma_\xi^\mathcal{E}(E(x) \otimes Z(x)AE(y) \otimes Z(y)) \\ &= \sum_{x, y} \Gamma_\xi^\mathcal{E}(P(x) \otimes Q(x)AP(y) \otimes Q(y)) = \Gamma_\xi^\mathcal{E}(\text{RAR}).\end{aligned}$$

In the final line, we have used the fact that $E(x) = P(x) + P^c(x)$ and $Z(x) = Q(x) + Q^c(x)$, together with (iv).

- (vi): We may write

$$\begin{aligned}\mathcal{I}_\chi^*(E(x)A) &= \Gamma_\xi^\mathcal{E}(\text{R}(E(x)A \otimes \mathbf{1}_A)\text{R}) \\ &= \Gamma_\xi^\mathcal{E}(P(x)AP(x) \otimes Q(x)) \\ &= \Gamma_\xi^\mathcal{E}(P(x)AP(x) \otimes \mathbf{1}_s) - \Gamma_\xi^\mathcal{E}(P(x)AP(x) \otimes Q(x)^\perp) \\ &= \Gamma_\xi^\mathcal{E}(P(x)AP(x) \otimes \mathbf{1}_s) = \mathcal{I}_\chi^*(P(x)AP(x)).\end{aligned}$$

The first line follows from (v), and the third line follows from the definition $Q(x)^\perp := \mathbf{1}_s - Q(x)$. The final line is obtained by (v) and noting that $(P(x)AP(x) \otimes Q(x)^\perp)\text{R} = P(x)AP(x) \otimes Q(x)^\perp Q(x) = \mathbf{0}$. The relation $\mathcal{I}_\chi^*(AE(x)) = \mathcal{I}_\chi^*(P(x)AP(x))$ holding for all A trivially follows from above and by observing that $\mathcal{I}_\chi^*(E(x)A^*)^* = \mathcal{I}_\chi^*(AE(x))$ and $\mathcal{I}_\chi^*(P(x)A^*P(x))^* = \mathcal{I}_\chi^*(P(x)AP(x))$.

Similarly, we may write

$$\begin{aligned}\Lambda^*(Z(x)B) &= \Gamma_\xi^\mathcal{E}(\text{R}(\mathbf{1}_s \otimes Z(x)B)\text{R}) \\ &= \Gamma_\xi^\mathcal{E}(P(x) \otimes Q(x)BQ(x)) \\ &= \Gamma_\xi^\mathcal{E}(\mathbf{1}_s \otimes Q(x)BQ(x)) - \Gamma_\xi^\mathcal{E}(P(x)^\perp \otimes Q(x)BQ(x)) \\ &= \Gamma_\xi^\mathcal{E}(\mathbf{1}_s \otimes Q(x)BQ(x)) = \Lambda^*(Q(x)BQ(x)).\end{aligned}$$

(vii): We may write

$$\begin{aligned}
\mathcal{I}_x^*(A) &= \Gamma_\xi^\mathcal{E}(\mathbf{R}(A \otimes \mathbf{Z}(x))\mathbf{R}) \\
&= \Gamma_\xi^\mathcal{E}(\mathbf{P}(x)A\mathbf{P}(x) \otimes \mathbf{Q}(x)) \\
&= \Gamma_\xi^\mathcal{E}(\mathbf{P}(x)A\mathbf{P}(x) \otimes \mathbf{Z}(x)) - \Gamma_\xi^\mathcal{E}(\mathbf{P}(x)A\mathbf{P}(x) \otimes \mathbf{Q}^c(x)) \\
&= \Gamma_\xi^\mathcal{E}(\mathbf{P}(x)A\mathbf{P}(x) \otimes \mathbf{Z}(x)) = \mathcal{I}_x^*(\mathbf{P}(x)A\mathbf{P}(x)).
\end{aligned}$$

In the first line we have used (v), in the third line we use $\mathbf{Q}(x) = \mathbf{Z}(x) - \mathbf{Q}^c(x)$, and in the final line we use (iv). □

Let us highlight one interesting property of repeatable instruments: if \mathcal{I} is repeatable, then for all input states ρ , the output states will be perfectly distinguishable. For the input state ρ , we define the normalised post-measurement states as $\rho_x := \mathcal{I}_x(\rho)/\text{tr}[\mathcal{I}_x(\rho)]$ for any x satisfying $\text{tr}[\mathcal{I}_x(\rho)] > 0$. By item (vii) of the above proposition, the Schrödinger picture operations of a repeatable instrument satisfy $\mathcal{I}_x(T) = \mathbf{P}(x)\mathcal{I}_x(T)\mathbf{P}(x)$ for all x and $T \in \mathcal{T}(\mathcal{H}_s)$, and so ρ_x will have support only in the eigenvalue-1 eigenspace of $\mathbf{E}(x)$. But by item (iii), such eigenvalue-1 eigenspaces are orthogonal, and so it holds that $\rho_x\rho_y = \rho_y\rho_x = \mathbf{0}$ for all $x \neq y$.

F Conservation laws

Recall that by Definition 1 a channel Φ conserves N on average if $\Phi^*(N) = N$, while by Definition 2 Φ fully conserves N if $\Phi^*(N^k) = N^k$ for all $k \in \mathbb{N}$. We shall now show that full conservation is in fact equivalent to just the first two moments being conserved, and that is also equivalent to “invariance” of the unitary group generated by N under the action of Φ^* , i.e., that $\Phi^*(e^{itN}) = e^{itN}$ for all $t \in \mathbb{R}$.

Proposition F.1. *Let $\Phi : \mathcal{T}(\mathcal{H}) \rightarrow \mathcal{T}(\mathcal{H})$ be a channel, and let $N \in \mathcal{L}_s(\mathcal{H})$ be a self-adjoint operator. The following statements are equivalent:*

- (i) Φ fully conserves N .
- (ii) $\Phi^*(N^k) = N^k$ for $k = 1, 2$.
- (iii) $\Phi^*(e^{itN}) = e^{itN}$ for all $t \in \mathbb{R}$.

Proof. (i) \implies (ii): This is trivial.

(ii) \implies (i): Assume that $\Phi^*(N^k) = N^k$ for $k = 1, 2$. For any $k \geq 2$, Corollary A.1 implies that $\Phi^*(N^{k+1}) = \Phi^*(N^k N) = \Phi^*(N^k)N$. The claim follows by induction.

(ii) \implies (iii): Define $f(t) := \Phi^*(e^{itN})$. Since N is bounded, e^{itN} is bounded and strongly continuous, and Φ^* is a channel, then $f(t)$ is infinitely differentiable. Now assume that $\Phi^*(N^k) = N^k$ for $k = 1, 2$. It follows that

$$\frac{d}{dt}f(t) = \mathbf{i}\Phi^*(Ne^{itN}) = \mathbf{i}N\Phi^*(e^{itN}) = \mathbf{i}Nf(t),$$

where the second equality follows from Corollary A.1. Indeed, by induction we obtain $(d^k/dt^k)f(t) = \mathbf{i}^k N^k f(t)$ for all $k \in \mathbb{N}$. Since $f(0) = \mathbf{1}$, by Taylor expansion around $t = 0$ we observe that

$$\Phi^*(e^{itN}) = \sum_{k=0}^{\infty} \frac{t^k}{k!} \left(\frac{d^k}{dt^k} f(t) \right) \Big|_{t=0} = \sum_{k=0}^{\infty} \frac{\mathbf{i}^k t^k}{k!} N^k = e^{itN}.$$

(iii) \implies (i): Assume that $\Phi^*(e^{itN}) = e^{itN}$ for all $t \in \mathbb{R}$. It follows that

$$\frac{d^k}{dt^k} \Phi^*(e^{itN}) = \frac{d^k}{dt^k} e^{itN} \implies \Phi^*(N^k e^{itN}) = N^k e^{itN}$$

holds for all $k \in \mathbb{N}$ and all t . Since $e^{itN} = \mathbf{1}$ when $t = 0$, it follows that $\Phi^*(N^k) = N^k$ for all k . □

A property that channels may enjoy is “covariance” under the action of a unitary group, i.e., that $\Phi^*(e^{itN} A e^{-itN}) = e^{itN} \Phi^*(A) e^{-itN}$ holds for all $t \in \mathbb{R}$ and $A \in \mathcal{L}(\mathcal{H})$. While a channel may be covariant while not invariant—for example, $\Phi^*(\cdot) = \text{tr}[\cdot]\mathbf{1}$ such that $[\omega, N] = \mathbf{0}$ is covariant but not invariant—we now show that invariance implies covariance.

Corollary F.1. *Let $\Phi : \mathcal{T}(\mathcal{H}) \rightarrow \mathcal{T}(\mathcal{H})$ be a channel, and let $N \in \mathcal{L}_s(\mathcal{H})$ be a self-adjoint operator. If Φ fully conserves N , then for all $t \in \mathbb{R}$ and $A \in \mathcal{L}(\mathcal{H})$ it holds that*

$$\Phi^*(e^{itN} A e^{-itN}) = e^{itN} \Phi^*(A) e^{-itN}$$

Proof. Let us define $V(t) := e^{itN}$ for notational simplicity. By Proposition F.1, full conservation of N by Φ implies that $\Phi^*(V(t)) = V(t)$ for all t . In particular, noting that $V(t)^* = V(-t)$, this implies that $\Phi^*(V(t)^* V(t)) - \Phi^*(V(t)^*) \Phi^*(V(t)) = \Phi^*(\mathbb{1}) - \mathbb{1} = \mathbb{O}$. As such, by Corollary A.1 it follows that

$$\Phi^*(V(t) A V(-t)) = V(t) \Phi^*(A V(-t)) = V(t) \Phi^*(A) V(-t).$$

□

We note that the condition $\Phi^*(N^k) = N^k$ for $k = 1, 2$ was taken as a potential definition of conservation *simpliciter* in Ref. [93]. However, the authors here conjectured that, in finite dimensions, the condition $\Phi^*(N^2) = N^2$ may be dropped, and that (in our formulation) both average and full conservation are equivalent. We shall now address this issue: by a simple counter-example, we shall show that average and full conservation are in fact not equivalent for general channels, even in finite dimensions.

Let us consider a system $\mathcal{H} \simeq \mathbb{C}^3$ with an orthonormal basis $\{|-1\rangle, |0\rangle, |1\rangle\}$. Now consider $N = \sum_n n |n\rangle\langle n| \equiv |1\rangle\langle 1| - |-1\rangle\langle -1|$, and a channel Φ^* defined by

$$\Phi^*(A) = \langle 1|A|1\rangle |1\rangle\langle 1| + \langle -1|A|-1\rangle |-1\rangle\langle -1| + \langle +|A|+\rangle |0\rangle\langle 0|,$$

to hold for all $A \in \mathcal{L}(\mathcal{H})$, where we define $|+\rangle := \frac{1}{\sqrt{2}}(|1\rangle + |-1\rangle)$. It is simple to verify that $\Phi^*(N) = N$, that is, Φ conserves N on average. However, $\Phi^*(N^2) = \mathbb{1} \neq N^2$, and so Φ does not fully conserve N . Since full conservation is equivalent to invariance, then it follows that Φ is also not invariant. Indeed, we can easily verify that Φ is not covariant either; for example, if we choose $A = |+\rangle\langle +|$, then it holds that

$$\begin{aligned} \Phi^*(e^{itN} |+\rangle\langle +| e^{-itN}) &= \frac{1}{2}(|1\rangle\langle 1| + |-1\rangle\langle -1|) + \cos(t)^2 |0\rangle\langle 0|, \\ e^{itN} \Phi^*(|+\rangle\langle +|) e^{-itN} &= \frac{1}{2}(|1\rangle\langle 1| + |-1\rangle\langle -1|) + |0\rangle\langle 0|, \end{aligned}$$

which coincide only when t is an integer multiple of π .

While the above discussion shows that average conservation is in general a weaker condition than full conservation, we shall now show that in the special case of unitary channels, average and full conservation are equivalent:

Lemma F.1. *Let $\Phi(\cdot) := U(\cdot)U^*$ be a unitary channel, with $U \in \mathcal{L}(\mathcal{H})$ a unitary operator, and let $N \in \mathcal{L}_s(\mathcal{H})$ be a self-adjoint operator. The following statements are equivalent:*

- (i) $[U, N] = \mathbb{O}$.
- (ii) Φ conserves N on average.
- (iii) Φ fully conserves N .

Proof. (i) \iff (ii), (i) \implies (iii), and (iii) \implies (ii) are trivial. To show (iii) \implies (i), let us first write

$$[U, N]^* [U, N] = \Phi^*(N^2) + N^2 - \Phi^*(N)N - N\Phi^*(N).$$

If Φ fully conserves N , then the right hand side vanishes. But since the left hand side is a positive operator, then it holds that $[U, N] = \mathbb{O}$. Finally, we shall show that (ii) \implies (iii). Since Φ is unitary, then $\Phi^*(A^*)\Phi^*(B) = U^* A^* U U^* B U = U^* A^* B U = \Phi^*(A^* B)$ holds for all $A, B \in \mathcal{L}(\mathcal{H})$. If $\Phi^*(N) = N$, it follows that $\Phi^*(N^2) = N^2$. The claim follows from Proposition F.1. □

G Bounds for measurement error under conservation laws

Here we provide quantitative trade-off relations for measurement error under additive conservation laws, both average and full.

Theorem G.1. Let $\mathcal{M} := (\mathcal{H}_A, \xi, \mathcal{E}, \mathcal{Z})$ be a measurement scheme for an observable acting in \mathcal{H}_S , and assume that \mathcal{E} conserves an additive quantity $N = N_S \otimes \mathbb{1}_A + \mathbb{1}_S \otimes N_A$ on average, where $N_S \in \mathcal{L}_s(\mathcal{H}_S)$ and $N_A \in \mathcal{L}_s(\mathcal{H}_A)$. Let $\|\epsilon(x)\|$ be the error in measuring the effects of the target observable \mathbf{E} , as defined in Eq. (7). Then for all $x \in \mathcal{X}$ it holds that

$$\|[\mathbf{E}(x), N_S] - \Lambda^*([\mathbf{Z}(x), N_A])\| \leq 2\|N_S\| \|\epsilon(x)\| + 2\|\Gamma_\xi^\mathcal{E}(N^2) - \Gamma_\xi^\mathcal{E}(N)^2\|^{\frac{1}{2}} \left(2\|\epsilon(x)\| + \|\mathbf{E}(x) - \mathbf{E}(x)^2\| \right)^{\frac{1}{2}}, \quad (35)$$

where $\Gamma_\xi^\mathcal{E}$ is the channel defined in Eq. (3), and Λ is the conjugate channel to $\mathcal{I}_\mathcal{X}$ defined in Eq. (5).

Proof. By Eq. (7), we have

$$[\mathbf{E}(x), N_S] - \Lambda^*([\mathbf{Z}(x), N_A]) = [N_S, \epsilon(x)] + [\Lambda^*(\mathbf{Z}(x)), N_S] - \Lambda^*([\mathbf{Z}(x), N_A]).$$

Since N is additive, average conservation of N by \mathcal{E} implies that $\Gamma_\xi^\mathcal{E}(N) = \Gamma_\xi(N) = N_S + \text{tr}[N_A \xi] \mathbb{1}_S$. It follows that for all $A \in \mathcal{L}(\mathcal{H}_S)$, $[A, N_S] = [A, \Gamma_\xi^\mathcal{E}(N)]$, and so we may write $[\Lambda^*(\mathbf{Z}(x)), N_S] = [\Gamma_\xi^\mathcal{E}(\mathbb{1}_S \otimes \mathbf{Z}(x)), \Gamma_\xi^\mathcal{E}(N)]$. Moreover, by additivity of N we have $\Gamma_\xi^\mathcal{E}([\mathbb{1}_S \otimes \mathbf{Z}(x), N]) = \Gamma_\xi^\mathcal{E}(\mathbb{1}_S \otimes [\mathbf{Z}(x), N_A]) = \Lambda^*([\mathbf{Z}(x), N_A])$. We may therefore write

$$[\mathbf{E}(x), N_S] - \Lambda^*([\mathbf{Z}(x), N_A]) = [N_S, \epsilon(x)] + [\Gamma_\xi^\mathcal{E}(\mathbb{1}_S \otimes \mathbf{Z}(x)), \Gamma_\xi^\mathcal{E}(N)] - \Gamma_\xi^\mathcal{E}([\mathbb{1}_S \otimes \mathbf{Z}(x), N]). \quad (36)$$

By the sesquilinear mapping $\langle\langle A|B \rangle\rangle = \Gamma_\xi^\mathcal{E}(A^*B) - \Gamma_\xi^\mathcal{E}(A^*)\Gamma_\xi^\mathcal{E}(B)$ and Corollary A.2 we obtain from Eq. (36) the bound

$$\|[\mathbf{E}(x), N_S] - \Lambda^*([\mathbf{Z}(x), N_A])\| \leq \| [N_S, \epsilon(x)] \| + 2\|\langle\langle N|N \rangle\rangle\|^{\frac{1}{2}} \|\langle\langle \mathbb{1}_S \otimes \mathbf{Z}(x) | \mathbb{1}_S \otimes \mathbf{Z}(x) \rangle\rangle\|^{\frac{1}{2}}.$$

Note that $\| [N_S, \epsilon(x)] \| \leq 2\|N_S\| \|\epsilon(x)\|$, and that $\langle\langle N|N \rangle\rangle = \Gamma_\xi^\mathcal{E}(N^2) - \Gamma_\xi^\mathcal{E}(N)^2$ by definition. Given that $\mathbb{1}_S \otimes \mathbf{Z}(x)$ and $\mathbf{E}(x)$ are effects and $\Gamma_\xi^\mathcal{E}$ is a channel, Lemma A.2 gives $\|\langle\langle \mathbb{1}_S \otimes \mathbf{Z}(x) | \mathbb{1}_S \otimes \mathbf{Z}(x) \rangle\rangle\| \leq 2\|\epsilon(x)\| + \|\mathbf{E}(x) - \mathbf{E}(x)^2\|$. We thus obtain the bound given in Eq. (35). \square

Proposition G.1. Let $\mathcal{M} := (\mathcal{H}_A, \xi, \mathcal{E}, \mathcal{Z})$ be a measurement scheme for an observable acting in \mathcal{H}_S , and assume that \mathcal{E} fully conserves an additive quantity $N = N_S \otimes \mathbb{1}_A + \mathbb{1}_S \otimes N_A$, where $N_S \in \mathcal{L}_s(\mathcal{H}_S)$ and $N_A \in \mathcal{L}_s(\mathcal{H}_A)$. Let $\|\epsilon(x)\|$ be the error in measuring the effects of the target observable \mathbf{E} , as defined in Eq. (7). Then for all $x \in \mathcal{X}$ it holds that

$$\|[\mathbf{E}(x), N_S] - \Lambda^*([\mathbf{Z}(x), N_A])\| \leq 2\|N_S\| \|\epsilon(x)\| + \frac{1}{2}\mathcal{Q}(N_A, \xi)^{\frac{1}{2}}, \quad (37)$$

where $\mathcal{Q}(N_A, \xi)$ denotes the quantum Fisher information of N_A in the state ξ . Additionally, if \mathbf{E} is an extremal observable and \mathcal{M} is a measurement scheme for \mathbf{E} , then for all $x \in \mathcal{X}$ it holds that

$$\|[\mathbf{E}(x), N_S] - \Lambda^*([\mathbf{Z}(x), N_A])\| \leq \mathcal{Q}(N_A, \xi)^{\frac{1}{2}} \|\mathbf{E}(x) - \mathbf{E}(x)^2\|^{\frac{1}{2}}. \quad (38)$$

Proof. Let $\{q_i, \phi_i\}$ be an arbitrary ensemble of unit vectors that satisfies $\xi = \sum_i q_i P_{\phi_i}$. We may thus write $\Gamma_\xi^\mathcal{E}(\cdot) = \sum_i q_i \Gamma_{\phi_i}^\mathcal{E}(\cdot)$, where $\Gamma_{\phi_i}^\mathcal{E}(\cdot) \equiv \Gamma_{P_{\phi_i}}^\mathcal{E}(\cdot)$. Given the additivity of N and the conservation law, we may rewrite Eq. (36) as

$$[\mathbf{E}(x), N_S] - \Lambda^*([\mathbf{Z}(x), N_A]) = [N_S, \epsilon(x)] + \sum_i q_i \left([\Gamma_{\phi_i}^\mathcal{E}(\mathbb{1}_S \otimes \mathbf{Z}(x)), \Gamma_{\phi_i}^\mathcal{E}(N)] - \Gamma_{\phi_i}^\mathcal{E}([\mathbb{1}_S \otimes \mathbf{Z}(x), N_A]) \right).$$

By the sesquilinear mappings $\langle\langle A|B \rangle\rangle_i := \Gamma_{\phi_i}^\mathcal{E}(A^*B) - \Gamma_{\phi_i}^\mathcal{E}(A^*)\Gamma_{\phi_i}^\mathcal{E}(B)$, Corollary A.2 and Lemma 3.1, we obtain the bounds

$$\|[\mathbf{E}(x), N_S] - \Lambda^*([\mathbf{Z}(x), N_A])\| \leq 2\|N_S\| \|\epsilon(x)\| + 2 \sum_i q_i \text{Var}(N_A, \phi_i)^{\frac{1}{2}} \|\Gamma_{\phi_i}^\mathcal{E}(\mathbb{1}_S \otimes \mathbf{Z}(x)^2) - \Gamma_{\phi_i}^\mathcal{E}(\mathbb{1}_S \otimes \mathbf{Z}(x))^2\|^{\frac{1}{2}}.$$

Since both $\mathbf{Z}(x)$ and $\Gamma_{\phi_i}^\mathcal{E}(\mathbb{1}_S \otimes \mathbf{Z}(x))$ are effects, we obtain

$$\|\Gamma_{\phi_i}^\mathcal{E}(\mathbb{1}_S \otimes \mathbf{Z}(x)^2) - \Gamma_{\phi_i}^\mathcal{E}(\mathbb{1}_S \otimes \mathbf{Z}(x))^2\|^{\frac{1}{2}} \leq \|\Gamma_{\phi_i}^\mathcal{E}(\mathbb{1}_S \otimes \mathbf{Z}(x)) - \Gamma_{\phi_i}^\mathcal{E}(\mathbb{1}_S \otimes \mathbf{Z}(x))^2\|^{\frac{1}{2}} \leq 1/2.$$

We thus arrive at the bound

$$\begin{aligned} \|\mathbb{E}(x), N_s] - \Lambda^*([Z(x), N_A])\| &\leq 2\|N_s\|\|\epsilon(x)\| + \sum_i q_i \text{Var}(N_A, \phi_i)^{\frac{1}{2}} \\ &\leq 2\|N_s\|\|\epsilon(x)\| + \left(\sum_i q_i \text{Var}(N_A, \phi_i)\right)^{\frac{1}{2}}, \end{aligned}$$

where the second line follows from the concavity of the square root. By choosing the ensemble $\{q_i, \phi_i\}$ that gives the quantum Fisher information as in Eq. (13), we arrive at the bound in Eq. (37).

Now assume that \mathbb{E} is an extremal observable [94]. This implies that for any pair of observables $\mathbb{E}^{(1)}$ and $\mathbb{E}^{(2)}$, and any $\lambda \in (0, 1)$, the effects of \mathbb{E} can be decomposed as $\mathbb{E}(x) = \lambda \mathbb{E}^{(1)}(x) + (1-\lambda) \mathbb{E}^{(2)}(x)$ only if $\mathbb{E} = \mathbb{E}^{(1)} = \mathbb{E}^{(2)}$. It follows that if \mathcal{M} is a measurement scheme for \mathbb{E} , that is, if $\epsilon = 0$, then $\Gamma_{\phi_i}^{\mathcal{E}}(\mathbb{1}_S \otimes Z(x)) = \mathbb{E}(x)$ for all i . Consequently, we obtain the bounds $\|\Gamma_{\phi_i}^{\mathcal{E}}(\mathbb{1}_S \otimes Z(x)^2) - \Gamma_{\phi_i}^{\mathcal{E}}(\mathbb{1}_S \otimes Z(x))^2\|^{\frac{1}{2}} \leq \|\mathbb{E}(x) - \mathbb{E}(x)^2\|^{\frac{1}{2}}$ for all i , which gives

$$\|\mathbb{E}(x), N_s] - \Lambda^*([Z(x), N_A])\| \leq 2 \left(\sum_i q_i \text{Var}(N_A, \phi_i)\right)^{\frac{1}{2}} \|\mathbb{E}(x) - \mathbb{E}(x)^2\|^{\frac{1}{2}}.$$

Once again choosing the ensemble that gives the quantum Fisher information, we arrive at Eq. (38). \square

H Bounds for measurement disturbance under conservation laws

Here we provide quantitative trade-off relations for measurement disturbance under additive conservation laws, both average and full. Note that here, the observable that may or may not be disturbed is not necessarily the same observable that is measured by the instrument \mathcal{I} .

Theorem H.1. *Let $\mathcal{M} := (\mathcal{H}_A, \xi, \mathcal{E}, Z)$ be a measurement scheme for an instrument \mathcal{I} acting in \mathcal{H}_S , and assume that \mathcal{E} conserves an additive quantity $N = N_S \otimes \mathbb{1}_A + \mathbb{1}_S \otimes N_A$ on average, where $N_S \in \mathcal{L}_s(\mathcal{H}_S)$ and $N_A \in \mathcal{L}_s(\mathcal{H}_A)$. Let $\|\delta(y)\|$ be the disturbance of the effects of an observable $F := \{F(y) : y \in \mathcal{Y}\}$ caused by \mathcal{I} , as defined in Eq. (8). Then for all $y \in \mathcal{Y}$ it holds that*

$$\|[F(y), N_s] - \mathcal{I}_{\mathcal{X}}^*([F(y), N_s])\| \leq 2\|N_s\|\|\delta(y)\| + 2\|\Gamma_{\xi}^{\mathcal{E}}(N^2) - \Gamma_{\xi}^{\mathcal{E}}(N)^2\|^{\frac{1}{2}} \|\mathcal{I}_{\mathcal{X}}^*(F(y)^2) - \mathcal{I}_{\mathcal{X}}^*(F(y))^2\|^{\frac{1}{2}}, \quad (39)$$

where $\Gamma_{\xi}^{\mathcal{E}}$ is the channel defined in Eq. (3). If F is non-disturbed by \mathcal{I} , that is, if $\delta = 0$, then for all $y \in \mathcal{Y}$ it holds that

$$\|[F(y), N_s] - \mathcal{I}_{\mathcal{X}}^*([F(y), N_s])\| \leq 2\|\Gamma_{\xi}^{\mathcal{E}}(N^2) - \Gamma_{\xi}^{\mathcal{E}}(N)^2\|^{\frac{1}{2}} \|\mathcal{I}_{\mathcal{X}}^*(F(y)^2) - F(y)^2\|^{\frac{1}{2}}. \quad (40)$$

Proof. By Eq. (8), we may write

$$[F(y), N_s] - \mathcal{I}_{\mathcal{X}}^*([F(y), N_s]) = [N_s, \delta(y)] + [\mathcal{I}_{\mathcal{X}}^*(F(y)), N_s] - \mathcal{I}_{\mathcal{X}}^*([F(y), N_s]).$$

Since N is additive, average conservation of N by \mathcal{E} implies that $\Gamma_{\xi}^{\mathcal{E}}(N) = \Gamma_{\xi}(N) = N_S + \text{tr}[N_A \xi] \mathbb{1}_S$. It follows that for all $A \in \mathcal{L}(\mathcal{H}_S)$, $[A, N_s] = [A, \Gamma_{\xi}^{\mathcal{E}}(N)]$. We may therefore write $[\mathcal{I}_{\mathcal{X}}^*(F(y)), N_s] = [\Gamma_{\xi}^{\mathcal{E}}(F(y) \otimes \mathbb{1}_A), \Gamma_{\xi}^{\mathcal{E}}(N)]$. Additionally, by additivity of N we may write $\Gamma_{\xi}^{\mathcal{E}}([F(y) \otimes \mathbb{1}_A, N]) = \Gamma_{\xi}^{\mathcal{E}}([F(y), N_s] \otimes \mathbb{1}_A) = \mathcal{I}_{\mathcal{X}}^*([F(y), N_s])$. We thus arrive at

$$[F(y), N_s] - \mathcal{I}_{\mathcal{X}}^*([F(y), N_s]) = [N_s, \delta(y)] + [\Gamma_{\xi}^{\mathcal{E}}(F(y) \otimes \mathbb{1}_A), \Gamma_{\xi}^{\mathcal{E}}(N)] - \Gamma_{\xi}^{\mathcal{E}}([F(y) \otimes \mathbb{1}_A, N]). \quad (41)$$

By the sesquilinear mapping $\langle\langle A|B \rangle\rangle := \Gamma_{\xi}^{\mathcal{E}}(A^*B) - \Gamma_{\xi}^{\mathcal{E}}(A^*)\Gamma_{\xi}^{\mathcal{E}}(B)$ and Corollary A.2 we obtain from Eq. (41) the bound

$$\|[F(y), N_s] - \mathcal{I}_{\mathcal{X}}^*([F(y), N_s])\| \leq \|[N_s, \delta(y)]\| + 2\|\langle\langle N|N \rangle\rangle\|^{\frac{1}{2}} \|\langle\langle F(y) \otimes \mathbb{1}_A | F(y) \otimes \mathbb{1}_A \rangle\rangle\|^{\frac{1}{2}}. \quad (42)$$

By definition, we have $\langle\langle F(y) \otimes \mathbb{1}_A | F(y) \otimes \mathbb{1}_A \rangle\rangle = \mathcal{I}_{\mathcal{X}}^*(F(y)^2) - \mathcal{I}_{\mathcal{X}}^*(F(y))^2$ and $\langle\langle N|N \rangle\rangle = \Gamma_{\xi}^{\mathcal{E}}(N^2) - \Gamma_{\xi}^{\mathcal{E}}(N)^2$. Noting that $\|[N_s, \delta(y)]\| \leq 2\|N_s\|\|\delta(y)\|$, we thus obtain from Eq. (42) the bound given in Eq. (39). If F is non-disturbed by \mathcal{I} , then $\|\delta(y)\| = 0$ and $\mathcal{I}_{\mathcal{X}}^*(F(y)^2) = F(y)^2$ for all y . We thus arrive at Eq. (40). \square

Corollary H.1. Consider the set-up of Theorem H.1. For all $y \in \mathcal{Y}$ it also holds that

$$\|[\mathbf{F}(y), N_s] - \mathcal{I}_{\mathcal{X}}^*([\mathbf{F}(y), N_s])\| \leq 2\|N_s\|\|\delta(y)\| + 2\|\Gamma_{\xi}^{\mathcal{E}}(N^2) - \Gamma_{\xi}^{\mathcal{E}}(N)^2\|^{\frac{1}{2}} \left(2\|\delta(y)\| + \|\mathbf{F}(y) - \mathbf{F}(y)^2\| \right)^{\frac{1}{2}}.$$

Proof. Since $\mathbf{F}(y)$ are effects and $\mathcal{I}_{\mathcal{X}}^*$ is a channel, then by Lemma A.2 we have $\|\mathcal{I}_{\mathcal{X}}^*(\mathbf{F}(y)^2) - \mathcal{I}_{\mathcal{X}}^*(\mathbf{F}(y))^2\| \leq 2\|\delta(y)\| + \|\mathbf{F}(y) - \mathbf{F}(y)^2\|$. The claim immediately follows from Eq. (39). \square

Corollary H.2. If \mathbf{E} commutes with N_s , and is measured by the Lüders instrument \mathcal{I}^L , then Theorem H.1 does not rule out non-disturbance of any observable \mathbf{F} that commutes with \mathbf{E} .

Proof. Since $\mathbf{E}' \subset \mathcal{F}(\mathcal{I}_{\mathcal{X}}^{L*})$ always holds, then $[\mathbf{E}, \mathbf{F}] = \mathbf{0}$ implies that $\mathbf{F} \subset \mathcal{F}(\mathcal{I}_{\mathcal{X}}^{L*})$. But $[\mathbf{E}, N_s] = \mathbf{0}$ implies that $\mathcal{I}_{\mathcal{X}}^{L*}(N_s^k) = N_s^k$ for $k = 1, 2$. By the multiplicability theorem (Corollary A.1), it holds that $\mathcal{I}_{\mathcal{X}}^{L*}(N_s A) = N_s \mathcal{I}_{\mathcal{X}}^{L*}(A)$ and $\mathcal{I}_{\mathcal{X}}^{L*}(A N_s) = \mathcal{I}_{\mathcal{X}}^{L*}(A) N_s$ for all $A \in \mathcal{L}(\mathcal{H}_s)$, and so we obtain $\mathcal{I}_{\mathcal{X}}^{L*}([\mathbf{F}(y), N_s]) = [\mathbf{F}(y), N_s]$. In such a case, the lower bound of Eq. (40) vanishes, and so no constraints are imposed on non-disturbance for \mathbf{F} . \square

Proposition H.1. Let $\mathcal{M} := (\mathcal{H}_A, \xi, \mathcal{E}, \mathbf{Z})$ be a measurement scheme for an instrument \mathcal{I} acting in \mathcal{H}_s , and assume that \mathcal{E} fully conserves an additive quantity $N = N_s \otimes \mathbf{1}_A + \mathbf{1}_s \otimes N_A$, where $N_s \in \mathcal{L}_s(\mathcal{H}_s)$ and $N_A \in \mathcal{L}_s(\mathcal{H}_A)$. Let $\|\delta(y)\|$ be the disturbance of the effects of an observable $\mathbf{F} = \{\mathbf{F}(y) : y \in \mathcal{Y}\}$ caused by \mathcal{I} , as defined in Eq. (8). Then for all $y \in \mathcal{Y}$ it holds that

$$\|[\mathbf{F}(y), N_s] - \mathcal{I}_{\mathcal{X}}^*([\mathbf{F}(y), N_s])\| \leq 2\|N_s\|\|\delta(y)\| + \frac{1}{2}\mathcal{Q}(N_A, \xi)^{\frac{1}{2}}, \quad (43)$$

where $\mathcal{Q}(N_A, \xi)$ denotes the quantum Fisher information of N_A in the state ξ . Additionally, if \mathcal{I} is an extremal instrument, then for all $y \in \mathcal{Y}$ it holds that

$$\|[\mathbf{F}(y), N_s] - \mathcal{I}_{\mathcal{X}}^*([\mathbf{F}(y), N_s])\| \leq 2\|N_s\|\|\delta(y)\| + \mathcal{Q}(N_A, \xi)^{\frac{1}{2}} \|\mathcal{I}_{\mathcal{X}}^*(\mathbf{F}(y)^2) - \mathcal{I}_{\mathcal{X}}^*(\mathbf{F}(y))^2\|^{\frac{1}{2}}. \quad (44)$$

Proof. Let $\{q_i, \phi_i\}$ be an arbitrary ensemble of unit vectors that satisfies $\xi = \sum_i q_i P_{\phi_i}$. We may thus write $\Gamma_{\xi}^{\mathcal{E}}(\cdot) = \sum_i q_i \Gamma_{\phi_i}^{\mathcal{E}}(\cdot)$, where $\Gamma_{\phi_i}^{\mathcal{E}}(\cdot) \equiv \Gamma_{P_{\phi_i}}^{\mathcal{E}}(\cdot)$. By the conservation law and additivity of N , we may therefore rewrite Eq. (41) as

$$[\mathbf{F}(y), N_s] - \mathcal{I}_{\mathcal{X}}^*([\mathbf{F}(y), N_s]) = [N_s, \delta(y)] + \sum_i q_i \left([\Gamma_{\phi_i}^{\mathcal{E}}(\mathbf{F}(y) \otimes \mathbf{1}_A), \Gamma_{\phi_i}^{\mathcal{E}}(N)] - \Gamma_{\phi_i}^{\mathcal{E}}([\mathbf{F}(y) \otimes \mathbf{1}_A, N]) \right),$$

which, by the sesquilinear mappings $\langle\langle A|B \rangle\rangle_i := \Gamma_{\phi_i}^{\mathcal{E}}(A^* B) - \Gamma_{\phi_i}^{\mathcal{E}}(A^*) \Gamma_{\phi_i}^{\mathcal{E}}(B)$, Corollary A.2, and Lemma 3.1 gives the bound

$$\|[\mathbf{F}(y), N_s] - \mathcal{I}_{\mathcal{X}}^*([\mathbf{F}(y), N_s])\| \leq 2\|N_s\|\|\delta(y)\| + 2 \sum_i q_i \text{Var}(N_A, \phi_i)^{\frac{1}{2}} \|\Gamma_{\phi_i}^{\mathcal{E}}(\mathbf{F}(y)^2 \otimes \mathbf{1}_A) - \Gamma_{\phi_i}^{\mathcal{E}}(\mathbf{F}(y) \otimes \mathbf{1}_A)^2\|^{\frac{1}{2}}.$$

Since both $\mathbf{F}(y)$ and $\Gamma_{\phi_i}^{\mathcal{E}}(\mathbf{F}(y) \otimes \mathbf{1}_A)$ are effects, we have

$$\|\Gamma_{\phi_i}^{\mathcal{E}}(\mathbf{F}(y)^2 \otimes \mathbf{1}_A) - \Gamma_{\phi_i}^{\mathcal{E}}(\mathbf{F}(y) \otimes \mathbf{1}_A)^2\|^{\frac{1}{2}} \leq \|\Gamma_{\phi_i}^{\mathcal{E}}(\mathbf{F}(y) \otimes \mathbf{1}_A) - \Gamma_{\phi_i}^{\mathcal{E}}(\mathbf{F}(y) \otimes \mathbf{1}_A)^2\|^{\frac{1}{2}} \leq \frac{1}{2}.$$

We thus arrive at the bound

$$\begin{aligned} \|[\mathbf{F}(y), N_s] - \mathcal{I}_{\mathcal{X}}^*([\mathbf{F}(y), N_s])\| &\leq 2\|N_s\|\|\delta(y)\| + \sum_i q_i \text{Var}(N_A, \phi_i)^{\frac{1}{2}} \\ &\leq 2\|N_s\|\|\delta(y)\| + \left(\sum_i q_i \text{Var}(N_A, \phi_i) \right)^{\frac{1}{2}}, \end{aligned}$$

where the second line follows from the concavity of the square root. By choosing the ensemble $\{q_i, \phi_i\}$ that gives the quantum Fisher information as in Eq. (13), we arrive at Eq. (43).

Now assume that \mathcal{I} is an extremal instrument [59]. This implies that for any pair of instruments $\mathcal{I}^{(1)}$ and $\mathcal{I}^{(2)}$, and any $\lambda \in (0, 1)$, the operations of \mathcal{I} can be decomposed as $\mathcal{I}_x(\cdot) = \lambda \mathcal{I}_x^{(1)}(\cdot) + (1 - \lambda) \mathcal{I}_x^{(2)}(\cdot)$ only if $\mathcal{I} = \mathcal{I}^{(1)} = \mathcal{I}^{(2)}$. It holds that $\Gamma_{\phi_i}^{\mathcal{E}}(\cdot \otimes \mathbf{1}_A) = \mathcal{I}_{\mathcal{X}}^*(\cdot)$ for all i , and so we obtain

$$\|[\mathbf{F}(y), N_s] - \mathcal{I}_{\mathcal{X}}^*([\mathbf{F}(y), N_s])\| \leq 2\|N_s\|\|\delta(y)\| + 2 \left(\sum_i q_i \text{Var}(N_A, \phi_i) \right)^{\frac{1}{2}} \|\mathcal{I}_{\mathcal{X}}^*(\mathbf{F}(y)^2) - \mathcal{I}_{\mathcal{X}}^*(\mathbf{F}(y))^2\|^{\frac{1}{2}}.$$

Once again, by choosing the ensemble that gives the quantum Fisher information, we arrive at Eq. (44). \square

I Proof for Generalised WAY theorem 2

Here, we shall provide a detailed proof for Theorem 3.3 presented in the main text.

Theorem I.1 (Generalised WAY theorem 2). *Let $\mathcal{M} := (\mathcal{H}_A, \xi, \mathcal{E}, Z)$ be a measurement scheme for an E-instrument \mathcal{I} acting in \mathcal{H}_S , and assume that \mathcal{E} conserves an additive quantity $N = N_S \otimes \mathbb{1}_A + \mathbb{1}_S \otimes N_A$ on average, where $N_S \in \mathcal{L}_s(\mathcal{H}_S)$ and $N_A \in \mathcal{L}_s(\mathcal{H}_A)$. If either \mathcal{I} is a measurement of the first kind, or the Yanase condition $[Z, N_A] = \mathbf{0}$ is satisfied, then for any effect $E(x)$ that has both eigenvalue 1 and 0, it holds that*

$$P[E(x), N_S]P = [E(x), PN_S P] = \mathbf{0}, \quad (45)$$

where $P := P_0(x) + P_1(x)$, with $P_0(x)$ and $P_1(x)$ orthogonal projections onto the eigenvalue-0 and eigenvalue-1 eigenspaces of $E(x)$, respectively.

Proof. Let us first note that $E(x)P = PE(x) = P_1(x)$, and $E(x)^\perp P = PE(x)^\perp = P_0(x)$, where $E(x)^\perp := \mathbb{1}_S - E(x)$. Indeed, $P_0(x)$ may equivalently be considered as the projection onto the eigenvalue-1 eigenspace of $E(x)^\perp$. Now define the operation $\Gamma_{\xi, P}^\mathcal{E} : \mathcal{L}(\mathcal{H}_S \otimes \mathcal{H}_A) \rightarrow \mathcal{L}(\mathcal{H}_S)$ as $\Gamma_{\xi, P}^\mathcal{E}(\cdot) := P\Gamma_\xi^\mathcal{E}(\cdot)P$, where $\Gamma_\xi^\mathcal{E}$ is the unital CP map defined in Eq. (3). If \mathcal{M} is a measurement scheme for E, then $\Gamma_{\xi, P}^\mathcal{E}(\mathbb{1}_S \otimes Z(x)) = P_1(x)$. Given $1 \geq \|Z(x)\| \geq \|\Gamma_{\xi, P}^\mathcal{E}(\mathbb{1}_S \otimes Z(x))\| = \|P_1(x)\| = 1$, it follows that $\|Z(x)\| = 1$. Therefore, by the two-positivity of CP maps, and the relation $A^*BA \leq \|B\|A^*A$ for self-adjoint B , we observe that

$$P_1(x) = \|Z(x)\|\Gamma_{\xi, P}^\mathcal{E}(\mathbb{1}_S \otimes Z(x)) \geq \Gamma_{\xi, P}^\mathcal{E}(\mathbb{1}_S \otimes Z(x)^2) \geq \Gamma_{\xi, P}^\mathcal{E}(\mathbb{1}_S \otimes Z(x))^2 = P_1(x),$$

and so

$$\Gamma_{\xi, P}^\mathcal{E}(\mathbb{1}_S \otimes Z(x)^2) = \Gamma_{\xi, P}^\mathcal{E}(\mathbb{1}_S \otimes Z(x))^2 = P_1(x). \quad (46)$$

Now note that by additivity of N , and the conservation law, it holds that $\Gamma_{\xi, P}^\mathcal{E}(N) = PN_S P + \text{tr}[N_A \xi]P$. If the Yanase condition holds, we may write

$$\begin{aligned} P[E(x), N_S]P &= [PE(x)P, PN_S P] \\ &= [\Gamma_{\xi, P}^\mathcal{E}(\mathbb{1}_S \otimes Z(x)), \Gamma_{\xi, P}^\mathcal{E}(N)] \\ &= \Gamma_{\xi, P}^\mathcal{E}([\mathbb{1}_S \otimes Z(x), N]) \\ &= \Gamma_{\xi, P}^\mathcal{E}(\mathbb{1}_S \otimes [Z(x), N_A]) = \mathbf{0}. \end{aligned}$$

The third line follows from Eq. (46) and the multiplicability theorem (Corollary A.1), and the final line follows from the Yanase condition. As such, we arrive at Eq. (45).

Now let us abandon the Yanase condition, but instead assume that \mathcal{I} is a first-kind measurement for E. This implies that $\Gamma_{\xi, P}^\mathcal{E}(E(x) \otimes \mathbb{1}_A) = P_1(x)$. Since $\|E(x)\| = 1$, then by the two-positivity of CP maps, and the relation $A^*BA \leq \|B\|A^*A$ for self-adjoint B , we obtain

$$P_1(x) = \|E(x)\|\Gamma_{\xi, P}^\mathcal{E}(E(x) \otimes \mathbb{1}_A) \geq \Gamma_{\xi, P}^\mathcal{E}(E(x)^2 \otimes \mathbb{1}_A) \geq \Gamma_{\xi, P}^\mathcal{E}(E(x) \otimes \mathbb{1}_A)^2 = P_1(x),$$

and so

$$\Gamma_{\xi, P}^\mathcal{E}(E(x)^2 \otimes \mathbb{1}_A) = \Gamma_{\xi, P}^\mathcal{E}(E(x) \otimes \mathbb{1}_A)^2 = P_1(x). \quad (47)$$

By the same arguments as in item (i) of Proposition E.1, one can show from Eq. (46) and Eq. (47) that $\Gamma_{\xi, P}^\mathcal{E}(\mathbb{1}_S \otimes Z(x)^n) = \Gamma_{\xi, P}^\mathcal{E}(E(x)^n \otimes \mathbb{1}_A) = P_1(x)$ for all $n \in \mathbb{N}$. Consequently, by the same arguments as in item (iv) of Proposition E.1, it follows that $\Gamma_{\xi, P}^\mathcal{E}(\mathbb{1}_S \otimes (Z(x) - Q(x))) = \mathbf{0}$, where $Q(x)$ is the projection onto the eigenvalue-1 eigenspace of $Z(x)$, and $\Gamma_{\xi, P}^\mathcal{E}((E(x) - P_1(x)) \otimes \mathbb{1}_A) = \mathbf{0}$. Moreover, by Eq. (46), Eq. (47), the multiplicability theorem (Corollary A.1), and defining $Z(x)^\perp := \mathbb{1}_A - Z(x)$, it follows that

$$\begin{aligned} \Gamma_{\xi, P}^\mathcal{E}(E(x) \otimes Z(x)^\perp) &= \Gamma_{\xi, P}^\mathcal{E}(E(x) \otimes \mathbb{1}_A)\Gamma_{\xi, P}^\mathcal{E}(\mathbb{1}_S \otimes Z(x)^\perp) = P_1(x)PE(x)^\perp P = P_1(x)P_0(x) = \mathbf{0}, \\ \Gamma_{\xi, P}^\mathcal{E}(E(x)^\perp \otimes Z(x)) &= \Gamma_{\xi, P}^\mathcal{E}(E(x)^\perp \otimes \mathbb{1}_A)\Gamma_{\xi, P}^\mathcal{E}(\mathbb{1}_S \otimes Z(x)) = PE(x)^\perp PP_1(x) = P_0(x)P_1(x) = \mathbf{0}, \end{aligned}$$

and so by the same arguments as in items (v) and (vi) of Proposition E.1 it follows that

$$\Gamma_{\xi, P}^\mathcal{E}(E(x)A \otimes \mathbb{1}_A) = \Gamma_{\xi, P}^\mathcal{E}(AE(x) \otimes \mathbb{1}_A) = \Gamma_{\xi, P}^\mathcal{E}(P_1(x)AP_1(x) \otimes \mathbb{1}_A) \quad (48)$$

for all $A \in \mathcal{L}(\mathcal{H}_S)$. By additivity of N , and the conservation law, we may therefore write

$$\begin{aligned} \mathsf{P}[\mathsf{E}(x), N_S] \mathsf{P} &= [\mathsf{P}\mathsf{E}(x)\mathsf{P}, \mathsf{P}N_S\mathsf{P}] \\ &= [\Gamma_{\xi, \mathsf{P}}^{\mathcal{E}}(\mathsf{E}(x) \otimes \mathbb{1}_A), \Gamma_{\xi, \mathsf{P}}^{\mathcal{E}}(N)] \\ &= \Gamma_{\xi, \mathsf{P}}^{\mathcal{E}}([\mathsf{E}(x) \otimes \mathbb{1}_A, N]) \\ &= \Gamma_{\xi, \mathsf{P}}^{\mathcal{E}}([\mathsf{E}(x), N_S] \otimes \mathbb{1}_A) = \mathbb{O}. \end{aligned}$$

The third line follows from Eq. (47) and Corollary A.1, while the final line follows from Eq. (48). Once again we arrive at Eq. (45). \square

J The Weak Yanase condition from conservation laws

Thus far, we have only considered the case where the measurement interaction \mathcal{E} between system and apparatus conserves an additive quantity N . However, pointer objectification will also result in state changes, and it may be the case that the expected value of N will change as a result. Now let us provide a generalised prescription of measurement schemes that captures also the state changes due to pointer objectification. Recall that $\mathcal{M} := (\mathcal{H}_A, \xi, \mathcal{E}, \mathsf{Z})$ is a measurement scheme for an observable E acting in \mathcal{H}_S if $\mathsf{E}(x) = \Gamma_{\xi} \circ \mathcal{E}^*(\mathbb{1}_S \otimes \mathsf{Z}(x))$. Now consider the tuple $\tilde{\mathcal{M}} := (\mathcal{H}_A, \xi, \mathcal{J})$, where $\mathcal{J} := \{\mathcal{J}_x : x \in \mathcal{X}\}$ is an instrument acting in $\mathcal{H}_S \otimes \mathcal{H}_A$. $\tilde{\mathcal{M}}$ is also a measurement scheme for E if $\mathsf{E}(x) = \Gamma_{\xi} \circ \mathcal{J}_x^*(\mathbb{1}_S \otimes \mathbb{1}_A)$. It is straightforward to show that this is satisfied if \mathcal{J} is compatible with the ‘‘Heisenberg-evolved’’ pointer observable

$$\mathsf{Z}^{\tau}(x) := \mathcal{E}^*(\mathbb{1}_S \otimes \mathsf{Z}(x)), \quad (49)$$

that is, if $\mathcal{J}_x^*(\mathbb{1}_S \otimes \mathbb{1}_A) = \mathsf{Z}^{\tau}(x)$. We say that $\tilde{\mathcal{M}}$ obeys a full (average) conservation law if the channel $\mathcal{J}_{\mathcal{X}}$ fully (on average) conserves a quantity N . The operations \mathcal{J}_x can be constructed as a sequential application of the channel \mathcal{E} followed by the operations of some Z -compatible instrument acting in \mathcal{H}_A , the latter of which provides a physical characterisation of the pointer objectification process. In such a case, a sufficient condition for conservation of N by $\mathcal{J}_{\mathcal{X}}$ is the conservation of N by both \mathcal{E} and the Z -channel. But it may be the case that \mathcal{E} fully conserves N while the Z -channel conserves N only on average, and vice versa. In such cases, the channel $\mathcal{J}_{\mathcal{X}}$ will conserve N only on average.

By Lemma C.1, it holds that if $\mathcal{J}_{\mathcal{X}}$ conserves N on average, and if either Z^{τ} is sharp or if $\mathcal{J}_{\mathcal{X}}$ also fully conserves N , then

$$[\mathsf{Z}^{\tau}, N] = \mathbb{O}. \quad (50)$$

This commutation relation is known as the weak Yanase condition [32]. We note that if \mathcal{E} conserves N on average and if either Z^{τ} is sharp or if \mathcal{E} also fully conserves N , then the Yanase condition implies the weak Yanase condition. First, let us assume that Z^{τ} is sharp. Since $\mathsf{Z}(x)$ is an effect then by two-positivity of CP maps we have $\mathsf{Z}^{\tau}(x) = \mathcal{E}^*(\mathbb{1}_S \otimes \mathsf{Z}(x)) \geq \mathcal{E}^*(\mathbb{1}_S \otimes \mathsf{Z}(x)^2) \geq \mathcal{E}^*(\mathbb{1}_S \otimes \mathsf{Z}(x))^2 = \mathsf{Z}^{\tau}(x)$, and so we have $\mathcal{E}^*(\mathbb{1}_S \otimes \mathsf{Z}(x)^2) = \mathcal{E}^*(\mathbb{1}_S \otimes \mathsf{Z}(x))^2$. On the other hand, if \mathcal{E} fully conserves N then $\mathcal{E}^*(N^2) = \mathcal{E}^*(N)^2 = N^2$. In either case, by Corollary A.1 we have

$$[\mathsf{Z}^{\tau}(x), N] = [\mathcal{E}^*(\mathbb{1}_S \otimes \mathsf{Z}(x)), \mathcal{E}^*(N)] = \mathcal{E}^*([\mathbb{1}_S \otimes \mathsf{Z}(x), N]) = \mathcal{E}^*(\mathbb{1}_S \otimes [\mathsf{Z}(x), N_A]),$$

and so if $[\mathsf{Z}(x), N_A] = \mathbb{O}$, then $[\mathsf{Z}^{\tau}(x), N] = \mathbb{O}$. Moreover, if $\mathcal{E}(\cdot) = U(\cdot)U^*$ is a unitary channel, and \mathcal{E} conserves N , then $[\mathsf{Z}^{\tau}(x), N] = U^*(\mathbb{1}_S \otimes [\mathsf{Z}(x), N_A])U$. In such a case the weak Yanase condition is equivalent to the Yanase condition; multiplying both sides of the equality $U^*(\mathbb{1}_S \otimes [\mathsf{Z}(x), N_A])U = \mathbb{O}$ by U from the left and by U^* from the right shows that $[\mathsf{Z}^{\tau}(x), N] = \mathbb{O} \iff [\mathsf{Z}(x), N_A] = \mathbb{O}$. However, in general it may be the case that the weak Yanase condition is satisfied but the Yanase condition is violated.

The following proposition shows that if the weak Yanase condition is satisfied, then the measurability part of the WAY theorem will hold. Moreover, we see that there are cases where a large coherence of the conserved quantity in the apparatus is necessary for good measurements even without a full conservation law—for example, if either the interaction channel \mathcal{E} or the Z -channel conserves N only on average, but Z^{τ} is sharp, in which case the weak Yanase condition is guaranteed to hold.

Proposition J.1. Let $\mathcal{M} := (\mathcal{H}_A, \xi, \mathcal{E}, Z)$ be a measurement scheme for an observable acting in \mathcal{H}_S , and let $\|\epsilon(x)\|$ be the error in measuring the effects of the target observable E . Assume that \mathcal{M} satisfies the weak Yanase condition $[Z^\tau, N] = \mathbf{O}$, where Z^τ is the Heisenberg-evolved pointer observable defined in Eq. (49) and $N = N_S \otimes \mathbb{1}_A + \mathbb{1}_S \otimes N_A$, where $N_S \in \mathcal{L}_s(\mathcal{H}_S)$ and $N_A \in \mathcal{L}_s(\mathcal{H}_A)$. Then for all $x \in \mathcal{X}$ it holds that

$$\|[\mathbf{E}(x), N_S]\| \leq 2\|N_S\| \|\epsilon(x)\| + 2\text{Var}(N_A, \xi)^{\frac{1}{2}} \left(2\|\epsilon(x)\| + \|\mathbf{E}(x) - E(x)^2\| \right)^{\frac{1}{2}}, \quad (51)$$

and

$$\|[\mathbf{E}(x), N_S]\| \leq 2\|N_S\| \|\epsilon(x)\| + \frac{1}{2}\mathcal{Q}(N_A, \xi)^{\frac{1}{2}}, \quad (52)$$

where $\text{Var}(N_A, \xi) := \text{tr}[N_A^2 \xi] - \text{tr}[N_A \xi]^2$ and $\mathcal{Q}(N_A, \xi)$ is the quantum Fisher information of N_A in ξ as defined in Eq. (13). Additionally, if \mathcal{M} is a measurement scheme for E , then for any effect $E(x)$ that has both eigenvalue 1 and 0, it holds that

$$P[\mathbf{E}(x), N_S]P = [\mathbf{E}(x), PN_S P] = \mathbf{O}, \quad (53)$$

where $P = P_0(x) + P_1(x)$, with $P_0(x)$ and $P_1(x)$ orthogonal projections onto the eigenvalue-0 and eigenvalue-1 eigenspaces of $E(x)$, respectively.

Proof. By Eq. (7), we may write $\epsilon(x) := \Gamma_\xi^\mathcal{E}(\mathbb{1}_S \otimes Z(x)) - E(x) \equiv \Gamma_\xi(Z^\tau(x)) - E(x)$. By additivity of N we have $\Gamma_\xi(N) = N_S + \text{tr}[N_A \xi] \mathbb{1}_S$, and so we may write

$$[\mathbf{E}(x), N_S] = [N_S, \epsilon(x)] + [\Gamma_\xi(Z^\tau(x)), \Gamma_\xi(N)].$$

Since Γ_ξ is a channel, and the weak Yanase condition $[Z^\tau, N] = \mathbf{O}$ holds, then by the sesquilinear map $\langle\langle A|B \rangle\rangle := \Gamma_\xi(A^*B) - \Gamma_\xi(A^*)\Gamma_\xi(B)$ and Corollary A.2 we obtain

$$\|[\mathbf{E}(x), N_S]\| \leq 2\|N_S\| \|\epsilon(x)\| + 2\|\Gamma_\xi(N^2) - \Gamma_\xi(N)^2\|^{\frac{1}{2}} \|\Gamma_\xi(Z^\tau(x)^2) - \Gamma_\xi(Z^\tau(x))^2\|^{\frac{1}{2}}.$$

As shown in Lemma 3.1, additivity of N implies that $\|\Gamma_\xi(N^2) - \Gamma_\xi(N)^2\| = \text{Var}(N_A, \xi)$. On the other hand, by Lemma A.2 we obtain $\|\Gamma_\xi(Z^\tau(x)^2) - \Gamma_\xi(Z^\tau(x))^2\| \leq 2\|\epsilon(x)\| + \|\mathbf{E}(x) - E(x)^2\|$. As such, we obtain the bound in Eq. (51). Eq. (52) and Eq. (53) are trivially obtained by adapting the arguments in Proposition G.1 and Theorem I.1 to the above, i.e., by replacing Z with Z^τ , and replacing the operations $\Gamma_{\phi_i}^\mathcal{E}$ and $\Gamma_{\xi, P}^\mathcal{E}$ with the operations Γ_{ϕ_i} and $\Gamma_{\xi, P}$, respectively. \square

K Proof of the ‘‘converse’’ WAY theorem

Here, we provide a proof for Proposition 4.1 presented in the main text.

Proposition K.1. Let $\mathcal{M} := (\mathcal{H}_A, \xi, \mathcal{E}, Z)$ be a measurement scheme for an instrument \mathcal{I} acting in \mathcal{H}_S . Assume that \mathcal{E} fully conserves an additive quantity $N = N_S \otimes \mathbb{1}_A + \mathbb{1}_S \otimes N_A$, where $N_S \in \mathcal{L}_s(\mathcal{H}_S)$ and $N_A \in \mathcal{L}_s(\mathcal{H}_A)$, and that $\mathcal{I}_\mathcal{X}$ fully conserves N_S . Define the eigenspace of \mathcal{H}_A that is involved during the measurement process as

$$\mathcal{H}_A(\text{meas}) := \bigcup_{\rho \in \mathcal{S}(\mathcal{H}_S)} \text{supp}(\Lambda(\rho)) \cup \text{supp}(\xi) \subseteq \mathcal{H}_A,$$

where Λ is the conjugate channel to $\mathcal{I}_\mathcal{X}$ defined in Eq. (5). Then either $\mathcal{H}_A(\text{meas})$ is contained within a single degenerate eigenspace of N_A , or $\text{Var}(N_A, \xi) := \text{tr}[N_A^2 \xi] - \text{tr}[N_A \xi]^2$ must be large. Additionally, if \mathcal{I} is an extremal instrument, and if $\mathcal{H}_A(\text{meas})$ is not contained within a single degenerate eigenspace of N_A , then the quantum Fisher information $\mathcal{Q}(N_A, \xi)$ as defined in Eq. (13) must be large.

Proof. Let us first observe that by Eq. (10), if \mathcal{E} conserves N on average, then $N_S \in \mathcal{F}(\mathcal{I}_\mathcal{X}^*)$ implies that irrespective of the apparatus preparation ξ , it holds that $\text{tr}[N_A \Lambda(\rho)] = \text{tr}[N_A \xi]$ for all $\rho \in \mathcal{S}(\mathcal{H}_S)$ or, equivalently, that $\Lambda^*(N_A) = \text{tr}[N_A \xi] \mathbb{1}_S$. That is, the expected value of the apparatus part of the conserved quantity does not change as a result of the measurement interaction. While average conservation does not imply that the variance must also stay the same, this implication can be shown to follow in the case of full conservation. By

Definition 2, full conservation of N by \mathcal{E} implies that $\Gamma_\xi(N^2) = \Gamma_\xi^\mathcal{E}(N^2)$, where $\Gamma_\xi^\mathcal{E}$ is the channel defined in Eq. (3). Given that $\Gamma_\xi^\mathcal{E}(\cdot \otimes \mathbb{1}_A) = \mathcal{I}_\chi^*(\cdot)$ and $\Gamma_\xi^\mathcal{E}(\mathbb{1}_S \otimes \cdot) = \Lambda^*(\cdot)$, we thus obtain

$$\begin{aligned} N_S^2 + 2\text{tr}[N_A\xi]N_S + \text{tr}[N_A^2\xi]\mathbb{1}_S &= \Gamma_\xi^\mathcal{E}(N_S^2 \otimes \mathbb{1}_A) + 2\Gamma_\xi^\mathcal{E}((N_S \otimes \mathbb{1}_A)(\mathbb{1}_S \otimes N_A)) + \Gamma_\xi^\mathcal{E}(\mathbb{1}_S \otimes N_A^2) \\ &= \Gamma_\xi^\mathcal{E}(N_S^2 \otimes \mathbb{1}_A) + 2\Gamma_\xi^\mathcal{E}(N_S \otimes \mathbb{1}_A)\Gamma_\xi^\mathcal{E}(\mathbb{1}_S \otimes N_A) + \Gamma_\xi^\mathcal{E}(\mathbb{1}_S \otimes N_A^2) \\ &= \mathcal{I}_\chi^*(N_S^2) + 2\mathcal{I}_\chi^*(N_S)\Lambda^*(N_A) + \Lambda^*(N_A^2) \\ &= N_S^2 + 2\text{tr}[N_A\xi]N_S + \Lambda^*(N_A^2). \end{aligned}$$

In the second line, we have used the fact that full conservation of N_S by \mathcal{I}_χ implies that $\Gamma_\xi^\mathcal{E}(N_S^2 \otimes \mathbb{1}_A) = \Gamma_\xi^\mathcal{E}(N_S \otimes \mathbb{1}_A)^2 = N_S^2$, and the multiplicability theorem (Corollary A.1). It follows that $\Lambda^*(N_A^k) = \text{tr}[N_A^k\xi]\mathbb{1}_S$ for $k = 1, 2$. As such, for any input state ρ of the system to be measured, it holds that

$$\text{Var}(N_A, \Lambda(\rho)) = \text{tr}[N_A^2\Lambda(\rho)] - \text{tr}[N_A\Lambda(\rho)]^2 = \text{tr}[N_A^2\xi] - \text{tr}[N_A\xi]^2 = \text{Var}(N_A, \xi).$$

Now assume that ξ is an eigenstate of N_A , i.e., that there exists $c \in \mathbb{R}$ such that $N_A\xi = c\xi$. Since such a condition is equivalent to a vanishing variance, then we see that $\text{Var}(N_A, \Lambda(\rho)) = \text{Var}(N_A, \xi) = 0$, and so $\Lambda(\rho)$ must also be eigenstates of N_A with the same eigenvalue c . In fact, in such a case it holds that $\Lambda^*(N_A^k) = \Lambda^*(N_A)^k = \text{tr}[N_A^k\xi]\mathbb{1}_S$ for $k = 1, 2$, which implies that $\text{tr}[N_A^k\Lambda(\rho)] = \text{tr}[N_A^k\xi] = c^k$ for all $k \in \mathbb{N}$ and $\rho \in \mathcal{S}(\mathcal{H}_S)$. In other words, if ξ is an eigenstate of N_A , then N_A must be “effectively” fully degenerate, i.e., $\mathcal{H}_A(\text{meas})$ must be contained within a single degenerate eigenspace of N_A . Therefore, if $\mathcal{H}_A(\text{meas})$ contains more than one degenerate eigenspace of N_A , the apparatus must be prepared in a state with a large uncertainty in N_A .

Now assume that \mathcal{I} is an extremal instrument, i.e., that for any $\lambda \in (0, 1)$, the operations of \mathcal{I} admit a decomposition $\mathcal{I}_x(\cdot) = \lambda\mathcal{I}_x^{(1)}(\cdot) + (1 - \lambda)\mathcal{I}_x^{(2)}(\cdot)$ only if $\mathcal{I} = \mathcal{I}^{(1)} = \mathcal{I}^{(2)}$. In such a case, it follows that if $\mathcal{M} := (\mathcal{H}_A, \xi, \mathcal{E}, \mathcal{Z})$ is a measurement scheme for \mathcal{I} , then for any pure state decomposition $\xi = \sum_i q_i P_{\phi_i}$, it holds that $(\mathcal{H}_A, \phi_i, \mathcal{E}, \mathcal{Z})$ must also be measurement schemes for \mathcal{I} . By the above arguments, it follows that unless N_A is effectively degenerate, then each ϕ_i must have a large uncertainty in N_A , i.e., the quantum Fisher information $\mathcal{Q}(N_A, \xi)$ must be large. \square

L Faithful fixed states and measurement disturbance

Recall from Proposition D.1 and Theorem H.1 that under a conservation law, an E-instrument \mathcal{I} will no disturb an observable F only if

$$\|[\mathbf{E}(x), F(y)]\| \leq 2\|\mathbf{E}(x) - \mathbf{E}(x)^2\|^{\frac{1}{2}}\|\mathcal{I}_\chi^*(F(y)^2) - F(y)^2\|^{\frac{1}{2}} \quad \forall x \in \mathcal{X}, y \in \mathcal{Y},$$

and

$$\|[\mathbf{F}(y), N_S] - \mathcal{I}_\chi^*([\mathbf{F}(y), N_S])\| \leq 2\|\Gamma_\xi^\mathcal{E}(N^2) - \Gamma_\xi^\mathcal{E}(N)^2\|^{\frac{1}{2}}\|\mathcal{I}_\chi^*(F(y)^2) - F(y)^2\|^{\frac{1}{2}} \quad \forall y \in \mathcal{Y}.$$

If $F \subset \mathcal{F}(\mathcal{I}_\chi^*) \implies F^2 \subset \mathcal{F}(\mathcal{I}_\chi^*)$, then the upper bound of the above equations vanish, and so F will be non-disturbed only if $[F, \mathbf{E}] = \mathbf{O}$ and $[\mathbf{F}(y), N_S] \in \mathcal{F}(\mathcal{I}_\chi^*)$. By the multiplicability theorem (Corollary A.1), the second condition can be shown to simplify to $[\mathbf{F}(y), N_S] = \mathcal{I}_\chi^*([\mathbf{F}(y), N_S]) = [\mathbf{F}(y), \mathcal{I}_\chi^*(N_S)]$, which reads $[\mathbf{F}(y), \mathcal{I}_\chi^*(N_S) - N_S] = \mathbf{O}$. In Appendix (D) we saw that the implication $F \subset \mathcal{F}(\mathcal{I}_\chi^*) \implies F^2 \subset \mathcal{F}(\mathcal{I}_\chi^*)$ holds if either F is sharp, rank-1, or a coarse-graining of a sharp observable. We now show that if $\mathcal{F}(\mathcal{I}_\chi^*)$ is a von Neumann algebra, so that $F \subset \mathcal{F}(\mathcal{I}_\chi^*) \implies F^2 \subset \mathcal{F}(\mathcal{I}_\chi^*)$, similar and stronger constraints will hold for all observables. To this end, let us first prove a useful lemma.

Lemma L.1. *Let $\mathcal{M} := (\mathcal{H}_A, \xi, \mathcal{E}, \mathcal{Z})$ be a measurement scheme for an E-instrument \mathcal{I} acting in \mathcal{H}_S . Assume that \mathcal{E} conserves an additive quantity $N = N_S \otimes \mathbb{1}_A + \mathbb{1}_S \otimes N_A$ on average, where $N_S \in \mathcal{L}_s(\mathcal{H}_S)$ and $N_A \in \mathcal{L}_s(\mathcal{H}_A)$, and that $\mathcal{F}(\mathcal{I}_\chi^*)$ is a von Neumann algebra. Then for all $A \in \mathcal{L}(\mathcal{H}_S)$, the following implication holds: $A \in \mathcal{F}(\mathcal{I}_\chi^*) \implies [A, N_S] \in \mathcal{F}(\mathcal{I}_\chi^*) \implies [A, N_S] \in \mathbf{E}'$.*

Proof. Recall that for all $A \in \mathcal{L}(\mathcal{H}_S)$, we have $\mathcal{I}_\chi^*(A) = \Gamma_\xi^\mathcal{E}(A \otimes \mathbb{1}_A)$, where $\Gamma_\xi^\mathcal{E}$ is the channel defined in Eq. (3). Average conservation of N by \mathcal{E} implies that

$$N_S + \text{tr}[N_A\xi]\mathbb{1}_S = \Gamma_\xi^\mathcal{E}(N_S \otimes \mathbb{1}_A) + \Gamma_\xi^\mathcal{E}(\mathbb{1}_S \otimes N_A).$$

Since $\mathcal{F}(\mathcal{I}_{\mathcal{X}}^*)$ is a von Neumann algebra, it follows that for all $A \in \mathcal{L}(\mathcal{H}_S)$, $A \in \mathcal{F}(\mathcal{I}_{\mathcal{X}}^*) \implies A^*A, AA^* \in \mathcal{F}(\mathcal{I}_{\mathcal{X}}^*)$ which, by Corollary A.1, implies that $A\Gamma_{\xi}^{\mathcal{E}}(B) = \Gamma_{\xi}^{\mathcal{E}}((A \otimes \mathbb{1}_A)B)$ and $\Gamma_{\xi}^{\mathcal{E}}(B)A = \Gamma_{\xi}^{\mathcal{E}}(B(A \otimes \mathbb{1}_A))$ for all $B \in \mathcal{L}(\mathcal{H}_S \otimes \mathcal{H}_A)$. Therefore, for all $A \in \mathcal{F}(\mathcal{I}_{\mathcal{X}}^*)$ we have

$$\begin{aligned} [A, N_S] &= [A, \Gamma_{\xi}^{\mathcal{E}}(N_S \otimes \mathbb{1}_A)] + [A, \Gamma_{\xi}^{\mathcal{E}}(\mathbb{1}_S \otimes N_A)] \\ &= \Gamma_{\xi}^{\mathcal{E}}([A \otimes \mathbb{1}_A, N_S \otimes \mathbb{1}_A]) + \Gamma_{\xi}^{\mathcal{E}}([A \otimes \mathbb{1}_A, \mathbb{1}_S \otimes N_A]) \\ &= \Gamma_{\xi}^{\mathcal{E}}([A, N_S] \otimes \mathbb{1}_A) = \mathcal{I}_{\mathcal{X}}^*([A, N_S]). \end{aligned}$$

Consequently, we see that $A \in \mathcal{F}(\mathcal{I}_{\mathcal{X}}^*) \implies [A, N_S] \in \mathcal{F}(\mathcal{I}_{\mathcal{X}}^*)$. But as shown in Lemma C.1, if $\mathcal{F}(\mathcal{I}_{\mathcal{X}}^*)$ is a von Neumann algebra then $\mathcal{F}(\mathcal{I}_{\mathcal{X}}^*) \subset \mathbb{E}'$. It follows that $A \in \mathcal{F}(\mathcal{I}_{\mathcal{X}}^*) \implies [A, N_S] \in \mathbb{E}'$. \square

We are now ready to prove the following:

Theorem L.1. *Let $\mathbb{E} := \{E(x) : x \in \mathcal{X}\}$ and $\mathbb{F} := \{F(y) : y \in \mathcal{Y}\}$ be observables acting in \mathcal{H}_S . Let $\mathcal{M} := (\mathcal{H}_A, \xi, \mathcal{E}, \mathcal{Z})$ be a measurement scheme for an \mathbb{E} -instrument \mathcal{I} , and assume that \mathcal{E} conserves an additive quantity $N = N_S \otimes \mathbb{1}_A + \mathbb{1}_S \otimes N_A$ on average, where $N_S \in \mathcal{L}_s(\mathcal{H}_S)$ and $N_A \in \mathcal{L}_s(\mathcal{H}_A)$. If $\mathcal{F}(\mathcal{I}_{\mathcal{X}}^*)$ is a von Neumann algebra, then the following hold:*

- (i) $\mathbb{F} \subset \mathcal{F}(\mathcal{I}_{\mathcal{X}}^*)$ only if \mathbb{F} commutes with \mathbb{E} , with $\Delta N_S := \mathcal{I}_{\mathcal{X}}^*(N_S) - N_S$, and with $\{[E(x), N_S] : x \in \mathcal{X}\}$.
- (ii) \mathcal{I} is a measurement of the first kind only if \mathbb{E} is a commutative observable that commutes with N_S .
- (iii) \mathcal{I} is repeatable only if \mathbb{E} is sharp and commutes with N_S .

Proof. (i): By Lemma L.1, $\mathcal{F}(\mathcal{I}_{\mathcal{X}}^*) \subset \mathbb{E}'$. Therefore, $\mathbb{F} \subset \mathcal{F}(\mathcal{I}_{\mathcal{X}}^*) \implies [\mathbb{E}, \mathbb{F}] = \mathbb{O}$. Moreover, the conservation law implies that $[\mathbb{F}, N_S] \subset \mathcal{F}(\mathcal{I}_{\mathcal{X}}^*)$ must hold. Since $\mathcal{F}(\mathcal{I}_{\mathcal{X}}^*)$ is an algebra, then $F(y) \in \mathcal{F}(\mathcal{I}_{\mathcal{X}}^*) \implies F(y)^2 \in \mathcal{F}(\mathcal{I}_{\mathcal{X}}^*)$. By Corollary A.1, it follows that $\mathcal{I}_{\mathcal{X}}^*(F(y)N_S) = F(y)\mathcal{I}_{\mathcal{X}}^*(N_S)$ and $\mathcal{I}_{\mathcal{X}}^*(N_SF(y)) = \mathcal{I}_{\mathcal{X}}^*(N_S)F(y)$, and so $\mathcal{I}_{\mathcal{X}}^*([F(y), N_S]) = [F(y), \mathcal{I}_{\mathcal{X}}^*(N_S)]$. As such, $\mathcal{I}_{\mathcal{X}}^*([F(y), N_S]) = [F(y), N_S] \implies [F(y), \mathcal{I}_{\mathcal{X}}^*(N_S) - N_S] = \mathbb{O}$. Moreover, by Lemma L.1 $\mathbb{F} \subset \mathcal{F}(\mathcal{I}_{\mathcal{X}}^*)$ implies $[F(y), N_S] \in \mathbb{E}'$. Since $[A, [B, C]] = [B, [A, C]] + [C, [B, A]]$ holds for all $A, B, C \in \mathcal{L}(\mathcal{H}_S)$, while $[\mathbb{E}, \mathbb{F}] = \mathbb{O}$, then this implies that $[F(y), [E(x), N_S]] = \mathbb{O}$.

(ii): \mathcal{I} is a measurement of the first kind only if $\mathbb{E} \subset \mathcal{F}(\mathcal{I}_{\mathcal{X}}^*)$. Commutativity of \mathbb{E} follows from (i). Now, let us define $N_S(t) := e^{itE(x)}N_S e^{-itE(x)}$. We may write

$$\frac{d}{dt}N_S(t) = i[E(x), N_S(t)],$$

from which we obtain

$$\begin{aligned} N_S(t) &= N_S + i \int_0^t dt_1 [E(x), N_S(t_1)] \\ &= N_S + i \int_0^t dt_1 [E(x), N_S] - \int_0^t dt_1 \int_0^{t_1} dt_2 e^{it_2 E(x)} [E(x), [E(x), N_S]] e^{-it_2 E(x)} \\ &= N_S + it[E(x), N_S]. \end{aligned}$$

In the final line we have used (i), which implies that $[E(x), [E(x), N_S]] = \mathbb{O}$ holds for all x . We thus obtain the inequality

$$2\|N_S\| \geq \|N_S(t) - N_S\| = |t| \| [E(x), N_S] \|$$

for all t . Given that N_S is a bounded operator, this is clearly satisfied only if $[E(x), N_S] = \mathbb{O}$.

- (iii): Since repeatability implies first-kindness, then by (ii) \mathbb{E} must commute with N_S . Sharpness of \mathbb{E} follows from item (i) of Proposition E.1 which gives $\mathcal{I}_{\mathcal{X}}^*(E(x)^2) = E(x)$ for a repeatable \mathcal{I} , and the fact that if $\mathcal{F}(\mathcal{I}_{\mathcal{X}}^*)$ is a von Neumann algebra then $\mathbb{E} \subset \mathcal{F}(\mathcal{I}_{\mathcal{X}}^*)$ implies that $\mathbb{E}^2 \subset \mathcal{F}(\mathcal{I}_{\mathcal{X}}^*)$. \square

We shall now give two examples where $\mathcal{F}(\mathcal{I}_{\mathcal{X}}^*)$ is a von Neumann algebra, and so the implications of Theorem L.1 hold.

Lemma L.2. *Consider the Lüders \mathbb{E} -instrument \mathcal{I}^L acting in \mathcal{H}_S , defined in Eq. (1). If either (i) $\dim(\mathcal{H}_S) < \infty$, or (ii) \mathbb{E} is commutative, then $\mathcal{F}(\mathcal{I}_{\mathcal{X}}^*)$ is a von Neumann algebra.*

Proof. Let us first consider (i). Define the complete mixture $\omega := \mathbb{1}_S / \dim(\mathcal{H}_S)$, which is faithful. It follows trivially that $\mathcal{I}_X^L(\omega) = \omega$, and so $\mathcal{F}(\mathcal{I}_X^L)$ contains a faithful state ω . By Lemma B.1 $\mathcal{F}(\mathcal{I}_X^L)$ is a von Neumann algebra. Now let us consider (ii). Recall that $\mathcal{F}(\mathcal{I}_X^*)$ is a von Neumann algebra if $\mathcal{F}(\mathcal{I}_X^*) = \{K_i, K_i^*\}'$, with $\{K_i\}$ any Kraus representation of \mathcal{I}_X [55]. But for a Lüders instrument, we have $\{K_i, K_i^*\}' = \{\sqrt{E(x)}\}' = E'$. While $E' \subset \mathcal{F}(\mathcal{I}_X^*)$ always holds, it was observed that in infinite-dimensional systems there exists E for which $\mathcal{F}(\mathcal{I}_X^*) \not\subset E'$ [78, 79]. However, it was shown in [63] that for binary observables, it always holds that $\mathcal{F}(\mathcal{I}_X^*) = E'$. Since binary observables are commutative, this led to the conjecture that the fixed-point set of the Lüders E -channel is the commutant of E for all commutative observables [78], making $\mathcal{F}(\mathcal{I}_X^*)$ a von Neumann algebra, which was later proven to be the case [76, 77]. \square

Let us highlight an interesting consequence of the above lemma:

Corollary L.1. *Let $\mathcal{M} := (\mathcal{H}_A, \xi, \mathcal{E}, \mathcal{Z})$ be a measurement scheme for an E -compatible Lüders instrument \mathcal{I}^L acting in \mathcal{H}_S . Assume that E is commutative, and that \mathcal{E} conserves an additive quantity $N = N_S \otimes \mathbb{1}_A + \mathbb{1}_S \otimes N_A$ on average. It holds that E commutes with N_S .*

Proof. If $E \subset E'$, then $\mathcal{F}(\mathcal{I}_X^*) = E'$ is a von Neumann algebra. Moreover, it holds that $E \subset \mathcal{F}(\mathcal{I}_X^*)$, so that the Lüders instrument for a commutative observable is a measurement of the first kind. It follows from item (ii) of Theorem L.1 that E must commute with N_S . \square

Lemma L.3. *Assume that $\dim(\mathcal{H}_S) < \infty$, and let $G := \{G(z) : z \in \mathcal{Z}\}$ be a rank-1 observable acting in \mathcal{H}_S . If an instrument \mathcal{I} does not disturb G , then $\mathcal{F}(\mathcal{I}_X^*)$ is a von Neumann algebra.*

Proof. Since $G(z)$ are rank-1 effects, we may write $G(z) = \lambda_z P_z$, with $\lambda_z \in (0, 1]$ and P_z a rank-1 projection. It follows that $G \subset \mathcal{F}(\mathcal{I}_X^*) \implies \{P_z\} \subset \mathcal{F}(\mathcal{I}_X^*)$. But we may write

$$\mathrm{tr}[P_z \mathcal{I}_X(P_z)] = \mathrm{tr}[\mathcal{I}_X^*(P_z) P_z] = \mathrm{tr}[P_z P_z] = \mathrm{tr}[P_z] = 1,$$

and so $G \subset \mathcal{F}(\mathcal{I}_X^*) \implies \{P_z\} \subset \mathcal{F}(\mathcal{I}_X)$. Consequently, we may construct the faithful state $\omega = \sum_z p_z P_z$ with $p_z > 0$ and $\sum_z p_z = 1$, so that $\omega \in \mathcal{F}(\mathcal{I}_X)$. By Lemma B.1, $\mathcal{F}(\mathcal{I}_X^*)$ is a von Neumann algebra. \square

M Non-faithful fixed states and measurement disturbance

In this section we analyse the structure of the fixed-point set of arbitrary channels, which need not contain a faithful state. From here, the results of the previous section are generalised. We then provide novel quantitative bounds for first-kind measurements which complement our generalisation of the WAY theorem given in Theorem 3.3.

Due to the Schauder–Tychonoff fixed point theorem [82], all channels $\Phi : \mathcal{T}(\mathcal{H}_S) \rightarrow \mathcal{T}(\mathcal{H}_S)$ have at least one fixed state. However, it may be that none of these are faithful. In such a case, the fixed-point set of the dual channel is not necessarily a von Neumann algebra, but rather forms an operator space [83]. This setting has been much less investigated, and its analysis forms the first part of this section. While the discussion thus far has been applicable for infinite-dimensional systems—except in some examples—in this section we shall always assume that $d := \dim(\mathcal{H}_S) < \infty$.

M.1 Fixed-point structure of arbitrary channels

Consider a channel $\Phi : \mathcal{T}(\mathcal{H}_S) \rightarrow \mathcal{T}(\mathcal{H}_S)$, and its dual in the Heisenberg picture $\Phi^* : \mathcal{L}(\mathcal{H}_S) \rightarrow \mathcal{L}(\mathcal{H}_S)$. We may define the channels

$$\Phi_{\mathrm{av}}(\cdot) := \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N \Phi^n(\cdot), \quad \Phi_{\mathrm{av}}^*(\cdot) := \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N \Phi^{*n}(\cdot), \quad (54)$$

where Φ^n denotes n consecutive applications of Φ . Note that these limits exist since $d < \infty$. According to the Jordan decomposition theorem, Φ^* is represented as a summation of projections onto eigenspaces multiplied by the corresponding eigenvalues, and nilpotent operators whose eigenspaces are invariant subspaces; Φ_{av}^* corresponds to the projection onto the subspace with eigenvalue 1. The fixed-point set $\mathcal{F}(\Phi^*)$ forms an operator space, i.e., a norm-closed vector subspace of the codomain of $\mathcal{F}(\Phi^*)$, and Φ_{av}^* is a CP projection onto $\mathcal{F}(\Phi^*)$.

Lemma M.1. Consider the channels Φ_{av}^* and Φ_{av} defined in Eq. (54). These have the following properties:

- (i) $\Phi^* \circ \Phi_{\text{av}}^* = \Phi_{\text{av}}^* \circ \Phi^* = \Phi_{\text{av}}^* \circ \Phi_{\text{av}}^* = \Phi_{\text{av}}^*$ and $\Phi \circ \Phi_{\text{av}} = \Phi_{\text{av}} \circ \Phi = \Phi_{\text{av}} \circ \Phi_{\text{av}} = \Phi_{\text{av}}$.
- (ii) $\Phi_{\text{av}}^*(\mathcal{L}(\mathcal{H}_S)) = \mathcal{F}(\Phi_{\text{av}}^*) = \mathcal{F}(\Phi^*)$ and $\Phi_{\text{av}}(\mathcal{T}(\mathcal{H}_S)) = \mathcal{F}(\Phi_{\text{av}}) = \mathcal{F}(\Phi)$.

Proof. (i) is trivial, and so we shall only prove (ii). Let us first consider the Heisenberg picture channel Φ_{av}^* . That $\mathcal{F}(\Phi^*) \subset \mathcal{F}(\Phi_{\text{av}}^*)$ is trivial. Conversely, for any $A \in \mathcal{F}(\Phi_{\text{av}}^*)$, by (i) we have $\Phi^*(A) = \Phi^* \circ \Phi_{\text{av}}^*(A) = \Phi_{\text{av}}^*(A) = A$, and therefore $\mathcal{F}(\Phi_{\text{av}}^*) \subset \mathcal{F}(\Phi^*)$. It follows that $\mathcal{F}(\Phi_{\text{av}}^*) = \mathcal{F}(\Phi^*)$. Similarly, for all $A \in \mathcal{L}(\mathcal{H}_S)$ it holds that $\Phi_{\text{av}}^* \circ \Phi_{\text{av}}^*(A) = \Phi_{\text{av}}^*(A)$, and thus $\Phi_{\text{av}}^*(\mathcal{L}(\mathcal{H}_S)) \subset \mathcal{F}(\Phi_{\text{av}}^*)$. That $\mathcal{F}(\Phi_{\text{av}}^*) \subset \Phi_{\text{av}}^*(\mathcal{L}(\mathcal{H}_S))$ is trivial, and so we also have $\Phi_{\text{av}}^*(\mathcal{L}(\mathcal{H}_S)) = \mathcal{F}(\Phi_{\text{av}}^*)$. The relations in (ii) for the Schrödinger picture channel Φ_{av} follow from similar arguments. \square

Now consider the state

$$\rho_0 := \Phi_{\text{av}} \left(\frac{1}{d} \mathbb{1}_S \right). \quad (55)$$

By Lemma M.1, it holds that $\rho_0 \in \mathcal{F}(\Phi_{\text{av}}) = \mathcal{F}(\Phi)$. We define by P the minimal support projection on ρ_0 :

$$P := \min\{Q : Q \text{ is a projection, } \rho_0 = Q\rho_0Q\}. \quad (56)$$

In other words, for all projections Q such that $\rho_0 = Q\rho_0Q$, it holds that $Q \geq P$. Note that if $P = \mathbb{1}_S$ then $\mathcal{F}(\Phi)$ contains a faithful state. The following lemma provides some useful properties of P .

Lemma M.2. Consider the state ρ_0 defined in Eq. (55), with the minimal support projection P as defined in Eq. (56), and $P^\perp := \mathbb{1}_S - P$ its orthogonal complement. The following hold:

- (i) $\Phi_{\text{av}}^*(P) = \mathbb{1}_S$ and $\Phi_{\text{av}}^*(P^\perp) = \mathbb{0}$.
- (ii) For all $A \in \mathcal{L}(\mathcal{H}_S)$, $\Phi_{\text{av}}^*(A) = \Phi_{\text{av}}^*(PAP)$.
- (iii) $P = \min\{Q : Q \text{ is a projection, } \rho = Q\rho Q \ \forall \rho \in \mathcal{F}(\Phi)\}$.
- (iv) $P = \min\{Q : Q \text{ is a projection, } \Phi_{\text{av}}^*(Q) = \mathbb{1}_S\}$.
- (v) $\Phi^*(P) \geq P$ and $\Phi^*(P^\perp) \leq P^\perp$.

Proof. (i): Since Φ_{av}^* is a channel, and $\mathbb{0} < P \leq \mathbb{1}_S$, it follows that $\mathbb{0} \leq \Phi_{\text{av}}^*(P) \leq \mathbb{1}_S$. But by Eq. (55) $\text{tr}[\Phi_{\text{av}}^*(P)] = d \text{tr}[\rho_0 P] = d$, and so $\Phi_{\text{av}}^*(P) = \mathbb{1}_S$. It trivially follows that $\Phi_{\text{av}}^*(P^\perp) = \Phi_{\text{av}}^*(\mathbb{1}_S) - \Phi_{\text{av}}^*(P) = \mathbb{1}_S - \mathbb{1}_S = \mathbb{0}$.

(ii): Since P is positive, then by (i) and Lemma A.3 it holds that $\Phi_{\text{av}}^*(P^\perp A) = \Phi_{\text{av}}^*(AP^\perp) = \mathbb{0}$ for all A . The claim follows by noting that we may write $A = (P + P^\perp)A(P + P^\perp)$.

(iii): By Lemma M.1 and (ii), for all $A \in \mathcal{L}(\mathcal{H}_S)$ and $\rho \in \mathcal{F}(\Phi)$ we have $\text{tr}[A\rho] = \text{tr}[\Phi_{\text{av}}^*(A)\rho] = \text{tr}[\Phi_{\text{av}}^*(PAP)\rho] = \text{tr}[AP\rho P]$, and so $\rho \in \mathcal{F}(\Phi) \implies \rho = P\rho P$. Since P is the minimal support projection on $\rho_0 \in \mathcal{F}(\Phi)$, the claim follows.

(iv): By (i), $\Phi_{\text{av}}^*(P) = \mathbb{1}_S$ holds. Suppose another projection Q satisfies $\Phi_{\text{av}}^*(Q) = \mathbb{1}_S$. Then (ii) implies that $\Phi_{\text{av}}^*(PQP) = \Phi_{\text{av}}^*(Q) = \mathbb{1}_S$. As we have $1 = \text{tr}[(\mathbb{1}_S/d)\Phi_{\text{av}}^*(PQP)] = \text{tr}[\rho_0 PQP] = \text{tr}[\rho_0 Q]$, it follows that $Q \geq P$.

(v): Since P is the smallest projection satisfying $\rho_0 = P\rho_0P$, while $\text{tr}[\rho_0 \Phi^*(P)] = \text{tr}[\Phi(\rho_0)P] = \text{tr}[\rho_0 P] = 1$, it follows that $\Phi^*(P) \geq P$, and hence $\Phi^*(P^\perp) = \Phi^*(\mathbb{1}_S) - \Phi^*(P) \leq \mathbb{1}_S - P = P^\perp$. \square

Let us now define the operations

$$\Phi_{\text{av},P}^*(\cdot) := P\Phi_{\text{av}}^*(\cdot)P, \quad \Phi_P^*(\cdot) := P\Phi^*(\cdot)P. \quad (57)$$

Note that $\Phi_{\text{av},P}^* : \mathcal{L}(\mathcal{H}_S) \rightarrow \mathcal{L}(\mathcal{H}_S)$ is not necessarily unital, since $\Phi_{\text{av},P}^*(\mathbb{1}_S) = P \leq \mathbb{1}_S$. However, the restriction of $\Phi_{\text{av},P}^*$ to $\mathcal{L}(P\mathcal{H}_S) \rightarrow \mathcal{L}(P\mathcal{H}_S)$, which is also denoted by $\Phi_{\text{av},P}^*$, is unital and hence a channel, since P is the identity in $\mathcal{L}(P\mathcal{H}_S)$ and $\Phi_{\text{av},P}^*(P) = P$. The same holds for Φ_P^* .

Lemma M.3. Consider the operations defined in Eq. (57). The following hold:

- (i) $\Phi_P^*(A) = \Phi_P^*(PAP)$ for all $A \in \mathcal{L}(\mathcal{H}_S)$.
- (ii) $\Phi_{\text{av},P}^*(A) = \Phi_{\text{av},P}^*(PAP)$ for all $A \in \mathcal{L}(\mathcal{H}_S)$.
- (iii) $\Phi_{\text{av},P}^*$ is a completely positive projection $\mathcal{L}(\mathcal{H}_S) \rightarrow \mathcal{L}(P\mathcal{H}_S)$.

- (iv) $\mathcal{F}(\Phi_{\text{av},P}^*) = \Phi_{\text{av},P}^*(\mathcal{L}(P\mathcal{H}_S))$.
- (v) Φ_{av}^* is a bijection from $\mathcal{F}(\Phi_{\text{av},P}^*)$ to $\mathcal{F}(\Phi^*)$.
- (vi) The inverse of $\Phi_{\text{av}}^* : \mathcal{F}(\Phi_{\text{av},P}^*) \rightarrow \mathcal{F}(\Phi^*)$ is $\text{Ad}_P : A \mapsto PAP$.

Proof. (i): By item (v) of Lemma M.2, it holds that $\mathbb{O} \leq \Phi_P^*(P^\perp) = P\Phi^*(P^\perp)P \leq PP^\perp P = \mathbb{O}$, and so $\Phi_P^*(P^\perp) = \mathbb{O}$. By Lemma A.3, it holds that $\Phi_P^*(P^\perp A) = \Phi_P^*(AP^\perp) = \mathbb{O}$ for all $A \in \mathcal{L}(\mathcal{H}_S)$. Since $A = (P + P^\perp)A(P + P^\perp)$, the claim follows.

(ii): By item (ii) of Lemma M.2, $\Phi_{\text{av}}^*(A) = \Phi_{\text{av}}^*(PAP)$. The claim immediately follows.

(iii): By item (i) of Lemma M.1 and (ii) above, we have

$$\Phi_{\text{av},P}^* \circ \Phi_{\text{av},P}^*(A) = \Phi_{\text{av},P}^*(P\Phi_{\text{av}}^*(A)P) = \Phi_{\text{av},P}^* \circ \Phi_{\text{av}}^*(A) = P\Phi_{\text{av}}^* \circ \Phi_{\text{av}}^*(A)P = P\Phi_{\text{av}}^*(A)P = \Phi_{\text{av},P}^*(A).$$

(iv): By (iii), it follows that for any $A \in \mathcal{L}(P\mathcal{H}_S)$, $\Phi_{\text{av},P}^*(A) \in \mathcal{F}(\Phi_{\text{av},P}^*)$ and so $\Phi_{\text{av},P}^*(\mathcal{L}(P\mathcal{H}_S)) \subset \mathcal{F}(\Phi_{\text{av},P}^*)$. The converse is trivial.

(v): For all $A \in \mathcal{F}(\Phi^*) = \mathcal{F}(\Phi_{\text{av}}^*)$, there exists an operator $PAP \in \mathcal{F}(\Phi_{\text{av},P}^*)$ such that $\Phi_{\text{av}}^*(PAP) = \Phi_{\text{av}}^*(A) = A$. Therefore, Φ_{av}^* is surjective. Now assume that there exists $A \in \mathcal{F}(\Phi_{\text{av},P}^*)$ such that $\Phi_{\text{av}}^*(A) = \mathbb{O}$. This implies that $A = \Phi_{\text{av},P}^*(A) = P\Phi_{\text{av}}^*(A)P = \mathbb{O}$. Therefore, Φ_{av}^* is injective.

(vi): Follows from above. □

The above results have the following useful consequence:

Proposition M.1. Consider the operations $\Phi_{\text{av},P}^*$ and Φ_P^* defined in Eq. (57). It holds that

$$\mathcal{F}(\Phi_{\text{av},P}^*) = \mathcal{F}(\Phi_P^*) = P\mathcal{F}(\Phi_{\text{av}}^*)P = P\mathcal{F}(\Phi^*)P$$

is a von Neumann algebra in $\mathcal{L}(P\mathcal{H}_S)$.

Proof. Recall that the operation $\Phi_{\text{av},P}^* : \mathcal{L}(P\mathcal{H}_S) \rightarrow \mathcal{L}(P\mathcal{H}_S)$ is unital, where the unit in $\mathcal{L}(P\mathcal{H}_S)$ is P . Moreover, ρ_0 as defined in Eq. (55) is a faithful fixed point of $\Phi_{\text{av},P}^*$ in $\mathcal{T}(P\mathcal{H}_S)$. By Lemma B.1, the fixed-point set $\mathcal{F}(\Phi_{\text{av},P}^*)$ is a von Neumann algebra in $\mathcal{L}(P\mathcal{H}_S)$. Now we need only show that $\mathcal{F}(\Phi_{\text{av},P}^*) = \mathcal{F}(\Phi_P^*) = P\mathcal{F}(\Phi_{\text{av}}^*)P = P\mathcal{F}(\Phi^*)P$. That $P\mathcal{F}(\Phi_{\text{av}}^*)P = P\mathcal{F}(\Phi^*)P$ trivially follows from Lemma M.1, which gives $\mathcal{F}(\Phi^*) = \mathcal{F}(\Phi_{\text{av}}^*)$. That $\mathcal{F}(\Phi_{\text{av},P}^*) = P\mathcal{F}(\Phi^*)P$ follows from item (vi) of Lemma M.3, since the map $\text{Ad}_P : A \mapsto PAP$ is a bijection from $\mathcal{F}(\Phi^*)$ to $\mathcal{F}(\Phi_{\text{av},P}^*)$. To show that $\mathcal{F}(\Phi_{\text{av},P}^*) = \mathcal{F}(\Phi_P^*)$, let us first define the operation $\Phi_{P,\text{av}}^* : \mathcal{L}(\mathcal{H}_S) \rightarrow \mathcal{L}(P\mathcal{H}_S)$ as

$$\Phi_{P,\text{av}}^*(\cdot) := \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N \Phi_P^{*n}(\cdot).$$

But by item (i) of Lemma M.3, $\Phi_P^*(A) = \Phi_P^*(PAP)$, and so $\Phi_P^{*n}(A) = P\Phi^{*n}(A)P$ for all $A \in \mathcal{L}(\mathcal{H}_S)$ and $n \in \mathbb{N}$. It follows that

$$\Phi_{P,\text{av}}^*(\cdot) = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N P\Phi^{*n}(\cdot)P = P\Phi_{\text{av}}^*(\cdot)P =: \Phi_{\text{av},P}^*(\cdot).$$

Since $\Phi_{\text{av},P}^* = \Phi_{P,\text{av}}^*$, it trivially follows that $\Phi_P^* \circ \Phi_{\text{av},P}^* = \Phi_{\text{av},P}^* \circ \Phi_P^* = \Phi_{\text{av},P}^* \circ \Phi_{\text{av},P}^* = \Phi_{\text{av},P}^*$. Therefore, $\mathcal{F}(\Phi_{\text{av},P}^*) = \mathcal{F}(\Phi_P^*)$ can be shown by the same arguments as in Lemma M.1 □

M.2 Measurement disturbance revisited

We are now ready to address the question of measurement disturbance, generalising the observations of Theorem L.1. As before, let $\mathcal{I} := \{\mathcal{I}_x : x \in \mathcal{X}\}$ be an E-compatible instrument, with $\mathcal{I}_{\mathcal{X}}(\cdot) := \sum_x \mathcal{I}_x(\cdot)$ the corresponding E-channel. By Eq. (54) and Eq. (57) we define

$$\mathcal{I}_{\text{av}}^*(\cdot) := \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N \mathcal{I}_{\mathcal{X}}^{*n}(\cdot), \quad \mathcal{I}_{\text{av},P}^*(\cdot) := P\mathcal{I}_{\text{av}}^*(\cdot)P, \quad \mathcal{I}_P^*(\cdot) := P\mathcal{I}_{\mathcal{X}}^*(\cdot)P,$$

where as in Eq. (56), P is the minimal support projection of $\rho_0 := \mathcal{I}_{\text{av}}(\frac{1}{d}\mathbb{1}_S)$, which corresponds with the minimal projection on the support of $\mathcal{F}(\mathcal{I}_X)$. By Proposition M.1, $\mathcal{F}(\mathcal{I}_{\text{av},P}^*) = \mathcal{F}(\mathcal{I}_P^*) = P\mathcal{F}(\mathcal{I}_{\text{av}}^*)P = P\mathcal{F}(\mathcal{I}_X^*)P$ is a von Neumann algebra in $\mathcal{L}(P\mathcal{H}_S)$. We define by $PEP := \{PE(x)P : x \in \mathcal{X}\}$ the restriction of E to an observable in $P\mathcal{H}_S$, which satisfies $\sum_x PE(x)P = P$, and $(PEP)' := \{A \in \mathcal{L}(P\mathcal{H}_S) : [PE(x)P, A] = \mathbb{O} \forall x\}$ denotes the commutant of PEP in $\mathcal{L}(P\mathcal{H}_S)$. PFP is similarly defined.

Before generalising Theorem L.1 for the case where $\mathcal{F}(\mathcal{I}_X)$ may not contain any faithful states, and thus $\mathcal{F}(\mathcal{I}_X^*)$ may not necessarily be a von Neumann algebra, let us first prove a generalisation of Lemma L.1.

Lemma M.4. *Let $\mathcal{M} := (\mathcal{H}_A, \xi, \mathcal{E}, \mathcal{Z})$ be a measurement scheme for an E -instrument \mathcal{I} acting in \mathcal{H}_S , and let P be the minimal support projection on $\mathcal{F}(\mathcal{I}_X)$. It holds that $\mathcal{F}(\mathcal{I}_P^*) \subset (PEP)'$. Additionally, if \mathcal{E} conserves an additive quantity $N = N_S \otimes \mathbb{1}_A + \mathbb{1}_S \otimes N_A$ on average, where $N_S \in \mathcal{L}_s(\mathcal{H}_S)$ and $N_A \in \mathcal{L}_s(\mathcal{H}_A)$, then for all $A \in \mathcal{L}(P\mathcal{H}_S)$ the following implication holds: $A \in \mathcal{F}(\mathcal{I}_P^*) \implies [A, PN_S P] \in \mathcal{F}(\mathcal{I}_P^*) \implies [A, PN_S P] \in (PEP)'$.*

Proof. By Eq. (3), let us define the operation $\Gamma_{\xi,P}^{\mathcal{E}} : \mathcal{L}(\mathcal{H}_S \otimes \mathcal{H}_A) \rightarrow \mathcal{L}(\mathcal{H}_S), B \mapsto P\Gamma_{\xi}^{\mathcal{E}}(B)P$. It is easily verified that $\mathcal{I}_P^*(A) = \Gamma_{\xi,P}^{\mathcal{E}}(A \otimes \mathbb{1}_A)$ for all $A \in \mathcal{L}(P\mathcal{H}_S)$, and $PE(x)P = \Gamma_{\xi,P}^{\mathcal{E}}(\mathbb{1}_S \otimes Z(x))$. Note that by the same arguments as item (i) of Lemma M.3, it can easily be shown that $\Gamma_{\xi,P}^{\mathcal{E}}(B) = \Gamma_{\xi,P}^{\mathcal{E}}((P \otimes \mathbb{1}_A)B(P \otimes \mathbb{1}_A))$ holds for all $B \in \mathcal{L}(\mathcal{H}_S \otimes \mathcal{H}_A)$. It follows that $\Gamma_{\xi,P}^{\mathcal{E}}$ is unital when restricted to $\mathcal{L}(P\mathcal{H}_S \otimes \mathcal{H}_A) \rightarrow \mathcal{L}(P\mathcal{H}_S)$, and we may equivalently write $PE(x)P = \Gamma_{\xi,P}^{\mathcal{E}}(P \otimes Z(x))$.

By Proposition M.1, $\mathcal{F}(\mathcal{I}_P^*)$ is a von Neumann algebra in $\mathcal{L}(P\mathcal{H}_S)$, and so for all $A \in \mathcal{L}(P\mathcal{H}_S)$, if $A \in \mathcal{F}(\mathcal{I}_P^*)$, then $A^*A, AA^* \in \mathcal{F}(\mathcal{I}_P^*)$. By Corollary A.1 it holds that for all $A \in \mathcal{F}(\mathcal{I}_P^*)$ and $B \in \mathcal{L}(P\mathcal{H}_S \otimes \mathcal{H}_A)$ we have $A\Gamma_{\xi,P}^{\mathcal{E}}(B) = \Gamma_{\xi,P}^{\mathcal{E}}((A \otimes \mathbb{1}_A)B)$ and $\Gamma_{\xi,P}^{\mathcal{E}}(B)A = \Gamma_{\xi,P}^{\mathcal{E}}(B(A \otimes \mathbb{1}_A))$. It follows that for all $A \in \mathcal{F}(\mathcal{I}_P^*)$ and $x \in \mathcal{X}$ we have

$$[A, PE(x)P] = [A, \Gamma_{\xi,P}^{\mathcal{E}}(P \otimes Z(x))] = \Gamma_{\xi,P}^{\mathcal{E}}([A \otimes \mathbb{1}_A, P \otimes Z(x)]) = \mathbb{O},$$

and so $\mathcal{F}(\mathcal{I}_P^*) \subset (PEP)'$.

Now let us assume that \mathcal{E} conserves an additive quantity $N = N_S \otimes \mathbb{1}_A + \mathbb{1}_S \otimes N_A$ on average. This implies that $P\Gamma_{\xi}(N)P = \Gamma_{\xi,P}^{\mathcal{E}}(N) = \Gamma_{\xi,P}^{\mathcal{E}}((P \otimes \mathbb{1}_A)N(P \otimes \mathbb{1}_A))$, and so

$$PN_S P + \text{tr}[N_A \xi]P = \Gamma_{\xi,P}^{\mathcal{E}}(PN_S P \otimes \mathbb{1}_A) + \Gamma_{\xi,P}^{\mathcal{E}}(P \otimes N_A).$$

Since $\mathcal{F}(\mathcal{I}_P^*)$ is a von Neumann algebra in $\mathcal{L}(P\mathcal{H}_S)$, then by Corollary A.1 and the arguments above, it follows that for all $A \in \mathcal{F}(\mathcal{I}_P^*)$ we have

$$\begin{aligned} [A, PN_S P] &= [A, \Gamma_{\xi,P}^{\mathcal{E}}(PN_S P \otimes \mathbb{1}_A)] + [A, \Gamma_{\xi,P}^{\mathcal{E}}(P \otimes N_A)] \\ &= \Gamma_{\xi,P}^{\mathcal{E}}([A \otimes \mathbb{1}_A, PN_S P \otimes \mathbb{1}_A]) + \Gamma_{\xi,P}^{\mathcal{E}}([A \otimes \mathbb{1}_A, P \otimes N_A]) \\ &= \Gamma_{\xi,P}^{\mathcal{E}}([A, PN_S P] \otimes \mathbb{1}_A) = \mathcal{I}_P^*([A, PN_S P]). \end{aligned}$$

We thus have $A \in \mathcal{F}(\mathcal{I}_P^*) \implies [A, PN_S P] \in \mathcal{F}(\mathcal{I}_P^*)$, and since $\mathcal{F}(\mathcal{I}_P^*) \subset (PEP)'$, it follows that $[A, PN_S P] \in (PEP)'$. \square

We are now ready to generalise Theorem L.1.

Theorem M.1. *Let $E := \{E(x) : x \in \mathcal{X}\}$ and $F := \{F(y) : y \in \mathcal{Y}\}$ be observables acting in \mathcal{H}_S . Let $\mathcal{M} := (\mathcal{H}_A, \xi, \mathcal{E}, \mathcal{Z})$ be a measurement scheme for an E -instrument \mathcal{I} , and assume that \mathcal{E} conserves an additive quantity $N = N_S \otimes \mathbb{1}_A + \mathbb{1}_S \otimes N_A$ on average, where $N_S \in \mathcal{L}_s(\mathcal{H}_S)$ and $N_A \in \mathcal{L}_s(\mathcal{H}_A)$. If P is the minimal support projection on $\mathcal{F}(\mathcal{I}_X)$, then the following hold:*

- (i) $F \subset \mathcal{F}(\mathcal{I}_X^*)$ only if PFP commutes with PEP , with $P\Delta N_S P := P\mathcal{I}_X^*(N_S)P - PN_S P$, and with $\{[PE(x)P, PN_S P] : x \in \mathcal{X}\}$.
- (ii) \mathcal{I} is a measurement of the first kind only if PEP is commutative and commutes with $PN_S P$.
- (iii) \mathcal{I} is repeatable only if PEP is sharp and commutes with $PN_S P$.

Proof. (i): By Proposition M.1, $P\mathcal{F}(\mathcal{I}_X^*)P = \mathcal{F}(\mathcal{I}_P^*)$, and so $F \subset \mathcal{F}(\mathcal{I}_X^*)$ implies that $PFP \subset \mathcal{F}(\mathcal{I}_P^*)$. By Lemma M.4, it holds that PFP must commute with PEP , and that $[PF(y)P, PN_S P] = \mathcal{I}_P^*([PF(y)P, PN_S P])$. Given that $\mathcal{F}(\mathcal{I}_P^*)$ is a von Neumann algebra, we have $\mathcal{I}_P^*((PF(y)P)^2) = \mathcal{I}_P^*(PF(y)P)^2 = (PF(y)P)^2$. By Corollary A.1 it follows that $[PF(y)P, PN_S P] = [PF(y)P, \mathcal{I}_P^*(PN_S P)]$. By item (i) of Lemma M.3, it holds that $\mathcal{I}_P^*(PN_S P) = \mathcal{I}_P^*(N_S) = P\mathcal{I}_X^*(N_S)P$. Therefore, PFP must commute with $P\Delta N_S P$. Moreover, by Lemma M.4 non-disturbance implies that $[PF(y)P, PN_S P] \in (PEP)'$, and since PFP commutes with PEP , this implies that $[PF(y)P, [PE(x)P, PN_S P]] = \mathbb{O}$ must hold.

- (ii): If \mathcal{I} is a measurement of the first kind, $E \subset \mathcal{F}(\mathcal{I}_\mathcal{X}^*)$, it follows from Proposition M.1 that $PEP \subset \mathcal{F}(\mathcal{I}_P^*)$. By Lemma M.4 PEP must be commutative, and since $\mathcal{F}(\mathcal{I}_P^*)$ is a von Neumann algebra, then for all spectral projections $R \in \mathcal{L}(\mathcal{PH}_S)$ of $PE(x)P$ it must hold that $[R, PN_S P]$ commutes with PEP . But this implies that $[[R, PN_S P], R^\perp] = \mathbf{0}$, where we define $R^\perp := P - R$. As such, given that $PR = RP = R$, it holds that $[RN_S P - PN_S R, R^\perp] = RN_S R^\perp + R^\perp N_S R = \mathbf{0}$. Multiplying from the left by R , we thus have $RN_S R^\perp = RN_S(P - R) = \mathbf{0}$, and so $RN_S P = RN_S R$. Since the right hand side is self-adjoint, and $RN_S P = RPN_S P$, it follows that $[R, PN_S P] = \mathbf{0}$. Since this relation holds for all spectral projections R of all effects of PEP , it follows that PEP must commute with $PN_S P$.
- (iii): Commutativity of PEP with $PN_S P$ follows from (ii) and the fact that repeatability implies first-kindness. Sharpness of PEP follows from the fact that the fixed points of a repeatable instrument can have support only in the eigenvalue-1 eigenspaces of E , as shown in Proposition E.1. □

Note that if $P = \mathbb{1}_S$, implying that $\mathcal{F}(\mathcal{I}_\mathcal{X})$ contains a faithful state so that $\mathcal{F}(\mathcal{I}_\mathcal{X}^*)$ is a von Neumann algebra, then the above theorem reduces to Theorem L.1. Interestingly, in the case of qubits such an equivalence will always hold, even if $\mathcal{F}(\mathcal{I}_\mathcal{X})$ does not contain a faithful state. We demonstrate this by an alternative proof for Proposition 6 in Ref. [4].

Corollary M.1. *If $\dim(\mathcal{H}_S) = 2$, then Theorem M.1 reduces to Theorem L.1.*

Proof. Let us consider the minimal support projection P on the fixed-point set $\mathcal{F}(\mathcal{I}_\mathcal{X})$. As all channels must have at least one fixed state, then when $\dim(\mathcal{H}_S) = 2$ it holds that, for any instrument \mathcal{I} , either $P = \mathbb{1}_S$ or P is a rank-1 projection. If $P = \mathbb{1}_S$ then $\mathcal{F}(\mathcal{I}_\mathcal{X})$ contains a faithful state, so that by Lemma B.1 $\mathcal{F}(\mathcal{I}_\mathcal{X}^*)$ is a von Neumann algebra, and Theorem M.1 reduces to Theorem L.1. Now assume that P is a rank-1 projection, so that for all $A \in \mathcal{L}(\mathcal{H}_S)$ it holds that $PAP = \lambda P$, with some $\lambda \in \mathbb{C}$. Recall from Lemma M.1 that $\mathcal{F}(\mathcal{I}_\mathcal{X}^*) = \mathcal{F}(\mathcal{I}_{\text{av}}^*)$. By item (ii) of Lemma M.2, for all $A \in \mathcal{F}(\mathcal{I}_\mathcal{X}^*)$ it holds that $A = \mathcal{I}_{\text{av}}^*(A) = \mathcal{I}_{\text{av}}^*(PAP) = \lambda \mathcal{I}_{\text{av}}^*(P) = \lambda \mathbb{1}_S$, and so $\mathcal{F}(\mathcal{I}_\mathcal{X}^*)$ is a trivial von Neumann algebra containing only operators proportional to the identity. In this case, the only non-disturbed observables are trivial, and will clearly commute with all of $\mathcal{L}(\mathcal{H}_S)$. In such a case we may simply replace P with $\mathbb{1}_S$ in items (i)-(iii) of Theorem M.1, so that it reduces to Theorem L.1. □

M.3 Non-disturbance and distinguishability

Here, we present some novel results regarding the structure of non-disturbing measurements that go beyond those in the preceding sections, indicating an intimate relationship between non-disturbance and distinguishability. These results hold for general instruments, are independent of conservation laws, and do not explicitly depend on the support projection P on the fixed-point set of the specific measurement channel $\mathcal{I}_\mathcal{X}$. First, let us show that if an instrument \mathcal{I} does not disturb a non-trivial observable, then there exists a family of distinguishable states that remain distinguishable after a non-selective measurement by \mathcal{I} .

Proposition M.2 (Non-disturbance implies distinguishability). *Consider an instrument \mathcal{I} acting in \mathcal{H}_S , and assume that \mathcal{I} does not disturb a non-trivial observable F . Then there exists a norm-1 observable $G := \{G(z) : z \in \mathcal{Z}\}$ acting in \mathcal{H}_S that is non-disturbed by \mathcal{I} , so that for every family of states $\{\rho_z : z \in \mathcal{Z}\}$ that are perfectly distinguishable by a G measurement, $\{\mathcal{I}_\mathcal{X}(\rho_z) : z \in \mathcal{Z}\}$ remain perfectly distinguishable by a G measurement. Moreover, if $\mathcal{F}(\mathcal{I}_\mathcal{X})$ contains a faithful state, then G can be taken as a sharp observable.*

Proof. Suppose that a non-trivial observable $F := \{F(y) : y \in \mathcal{Y}\}$ is non-disturbed by \mathcal{I} , i.e., $F \subset \mathcal{F}(\mathcal{I}_\mathcal{X}^*) = \mathcal{F}(\mathcal{I}_{\text{av}}^*)$. Given that $\mathcal{F}(\mathcal{I}_{\text{av},P}^*) = P\mathcal{F}(\mathcal{I}_{\text{av}}^*)P$ (see Proposition M.1), this implies that $PF(y)P \in \mathcal{F}(\mathcal{I}_{\text{av},P}^*)$. That F is non-trivial implies that there must be a y for which $PF(y)P$ is not proportional to P . If that were the case every $PF(y)P$ could be written as $PF(y)P = c_y P$ with some $c_y \geq 0$, which would imply that $F(y) = \mathcal{I}_{\text{av}}^*(F(y)) = \mathcal{I}_{\text{av}}^*(PF(y)P) = c_y \mathcal{I}_{\text{av}}^*(P) = c_y \mathbb{1}_S$, and so F would be a trivial observable. Therefore $\mathcal{F}(\mathcal{I}_{\text{av},P}^*) = P\mathcal{F}(\mathcal{I}_{\text{av}}^*)P$ is a nontrivial von Neumann algebra in $\mathcal{L}(\mathcal{PH}_S)$, and there exists a family of projections $R := \{R(z) : z \in \mathcal{Z}\} \subset \mathcal{F}(\mathcal{I}_{\text{av},P}^*)$ satisfying $R(z)R(y) = \delta_{z,y}R(z)$ and $\sum_z R(z) = P$. We may consider R as a sharp observable acting in \mathcal{PH}_S .

Using R , we may define a (generally unsharp) observable G acting in \mathcal{H}_S by

$$G := \{G(z) = \mathcal{I}_{\text{av}}^*(R(z)) : z \in \mathcal{Z}\},$$

where $\sum_z G(z) = \mathcal{I}_{\text{av}}^*(P) = \mathbb{1}_S$. Given that $\mathcal{I}_\mathcal{X}^* \circ \mathcal{I}_{\text{av}}^* = \mathcal{I}_{\text{av}}^*$ (see Lemma M.1), it follows that $G \subset \mathcal{F}(\mathcal{I}_\mathcal{X}^*)$, i.e., G is non-disturbed by \mathcal{I} . Moreover, since $PG(z)P = \mathcal{I}_{\text{av},P}^*(R(z)) = R(z)$ holds, then each (non-zero) effect of G

has at least one eigenvector with eigenvalue 1, and so $\|G(z)\| = 1$. Moreover, if $\mathcal{F}(\mathcal{I}_{\mathcal{X}})$ contains a faithful state, then $P = \mathbb{1}_S$, and so we have $G(z) = PG(z)P = R(z)$, implying that $G \equiv R$ is sharp. Now let us note that the family of states $\{\rho_z\}$ are perfectly distinguishable given a measurement of G if and only if $\rho_z = P(z)\rho_zP(z)$, where $P(z) \geq R(z)$ projects onto the eigenvalue-1 eigenspace of $G(z)$. In such a case it trivially holds that $\text{tr}[G(y)\rho_z] = \delta_{z,y}$. But since $G \subset \mathcal{F}(\mathcal{I}_{\mathcal{X}}^*)$, we also have

$$\text{tr}[G(y)\mathcal{I}_{\mathcal{X}}(\rho_z)] = \text{tr}[\mathcal{I}_{\mathcal{X}}^*(G(y))\rho_z] = \text{tr}[G(y)\rho_z] = \delta_{z,y},$$

and so $\{\mathcal{I}_{\mathcal{X}}(\rho_z)\}$ continues to be perfectly distinguishable by a G measurement. \square

In the special case where \mathcal{I} is a measurement of the first kind, we may strengthen the above result as follows:

Theorem M.2. *Let \mathcal{I} be an instrument compatible with a non-trivial observable E acting in \mathcal{H}_S . If \mathcal{I} is a measurement of the first kind, then E is described by a classical post-processing of a norm-1 observable $G := \{G(z) : z \in \mathcal{Z}\}$ with properties given in Proposition M.2, i.e.,*

$$E(x) = \sum_z p(x|z)G(z), \quad (58)$$

where $\{p(x|z)\}$ is a family of non-negative numbers that satisfy $\sum_x p(x|z) = 1$ for each z .

Proof. Assume that the E -instrument \mathcal{I} is a measurement of the first kind, that is, $E \subset \mathcal{F}(\mathcal{I}_{\mathcal{X}}^*)$. It follows that $PEP \subset \mathcal{F}(\mathcal{I}_{\text{av},P}^*)$. In fact, we can show that $PEP \subset \mathcal{F}(\mathcal{I}_{\text{av},P}^*) \cap \mathcal{F}(\mathcal{I}_{\text{av},P}^*)'$. First, recall from Proposition M.2 that if E is non-trivial then there exists a family of projections $R := \{R(z) : z \in \mathcal{Z}\} \subset \mathcal{F}(\mathcal{I}_{\text{av},P}^*)$, satisfying $R(z)R(y) = \delta_{z,y}R(z)$ and $\sum_z R(z) = P$. If $PEP \not\subset \mathcal{F}(\mathcal{I}_{\text{av},P}^*) \cap \mathcal{F}(\mathcal{I}_{\text{av},P}^*)'$, then R can be chosen so that $[PE(x)P, R(z)] \neq \mathbb{O}$ for some x and z . But note that $R \subset \mathcal{F}(\mathcal{I}_{\text{av},P}^*) = P\mathcal{F}(\mathcal{I}_{\mathcal{X}}^*)P$ implies that the observable $\{PT_x^*(R(z))P, PT_x^*(\mathbb{1}_S - R(z))P : x \in \mathcal{X}\}$ is a joint measurement for PEP and the sharp observable $\{R(z), P - R(z)\}$. By compatibility, it follows that $[PE(x)P, R(z)] = \mathbb{O}$ must hold for all x and z , and so PEP must be contained in the Abelian algebra $\mathcal{F}(\mathcal{I}_{\text{av},P}^*) \cap \mathcal{F}(\mathcal{I}_{\text{av},P}^*)'$. Consequently, $R \subset \mathcal{F}(\mathcal{I}_{\text{av},P}^*) \cap \mathcal{F}(\mathcal{I}_{\text{av},P}^*)'$ can be chosen so as to simultaneously diagonalise all $PE(x)P$, that is, we may write $PE(x)P = \sum_z p(x|z)R(z)$. Recalling that $E(x) = \mathcal{I}_{\text{av}}^*(E(x)) = \mathcal{I}_{\text{av}}^*(PE(x)P)$, then defining the observable G by $G(z) = \mathcal{I}_{\text{av}}^*(R(z))$ gives us Eq. (58). As in Proposition M.2 it holds that $G \subset \mathcal{F}(\mathcal{I}_{\mathcal{X}}^*)$, G is a norm-1 observable, if $\mathcal{F}(\mathcal{I}_{\mathcal{X}})$ contains a faithful state then G is also sharp, and if $\{\rho_z\}$ are perfectly distinguishable by a G measurement then so are $\{\mathcal{I}_{\mathcal{X}}(\rho_z)\}$. \square

Finally, we present the following implication of the above theorem:

Corollary M.2. *Let \mathcal{I} be an instrument compatible with a non-trivial observable E acting in \mathcal{H}_S , and assume that \mathcal{I} is a measurement of the first kind. For any outcome x associated with a non-trivial effect $E(x)$, and for any pair of unit vectors $\psi, \phi \in \mathcal{H}_S$ satisfying $E(x)\psi = \|E(x)\|\psi$ and $(\mathbb{1}_S - E(x))\phi = \|\mathbb{1}_S - E(x)\|\phi$, respectively, it holds that ψ and ϕ are orthogonal, and that $F(\mathcal{I}_{\mathcal{X}}(P_\psi), \mathcal{I}_{\mathcal{X}}(P_\phi)) = 0$, where $F(\rho, \sigma) := \text{tr}[\sqrt{\sqrt{\rho}\sigma\sqrt{\rho}}]$ is the fidelity between states ρ and σ .*

Proof. For each outcome x , we may coarse-grain E into a binary observable $\{E(x), E(\bar{x}) := \mathbb{1}_S - E(x)\}$. By Theorem M.2, first-kindness implies that $E(x) = \sum_z p(x|z)G(z)$ and $E(\bar{x}) = \sum_z p(\bar{x}|z)G(z)$, with $p(\bar{x}|z) := 1 - p(x|z)$, where $\{G(z) : z \in \mathcal{Z}\}$ is a norm-1 observable with properties given in Proposition M.2, while $\{p(x|z)\}$ is a family of non-negative numbers satisfying $\sum_{x'} p(x'|z) = 1$ for each z . Now let us define $p_{\max}(x) := \max_z \{p(x|z)\}$ and $p_{\min}(x) := \min_z \{p(x|z)\}$ as the maximum and minimum values of the set $\{p(x|z)\}$. We may thus define the sets $Z_{\max} := \{z \in \mathcal{Z} : p(x|z) = p_{\max}(x)\}$ and $Z_{\min} := \{z \in \mathcal{Z} : p(x|z) = p_{\min}(x)\}$. Using such sets, we may define $G(Z_{\max}) := \sum_{z \in Z_{\max}} G(z)$ and $G(Z_{\min}) := \sum_{z \in Z_{\min}} G(z)$. Since G is norm-1, then we may also define $P(Z_{\max}) := \sum_{z \in Z_{\max}} P(z)$ and $P(Z_{\min}) := \sum_{z \in Z_{\min}} P(z)$, where $P(z)$ is the projection onto the eigenvalue-1 eigenspace of $G(z)$. Since $E(x)$ is assumed to be non-trivial, then it must hold that $Z_{\max} \cap Z_{\min} = \emptyset$. If this were not so, it would hold that all $p(x|z)$ are the same, in which case $E(x) \propto \mathbb{1}_S$. Consequently, $P(Z_{\max})P(Z_{\min}) = \mathbb{O}$.

Now let us note that

$$\|E(x)\| = \sup_{\|\psi\|=1} \langle \psi | E(x) \psi \rangle = \sup_{\|\psi\|=1} \sum_z p(x|z) \langle \psi | G(z) \psi \rangle.$$

We may now show that a unit vector ψ satisfies $E(x)\psi = \|E(x)\|\psi$ if and only if $P(Z_{\max})\psi = \psi$. Let us first prove the only if statement. For any unit vector ψ , it holds that $\langle \psi | E(x) \psi \rangle \leq p_{\max}(x)$, which follows from the fact that $p(x|z)$ are positive numbers and that $\{\langle \psi | G(z) \psi \rangle\}$ is a probability distribution, with the upper bound being saturated when $\langle \psi | G(Z_{\max}) \psi \rangle = 1$. But this in turn is satisfied only if $\langle \psi | P(Z_{\max}) \psi \rangle = 1$, in which case

$P(Z_{\max})\psi = \psi$. As such, it follows that $\|E(x)\| = p_{\max}(x)$, and the unit vector ψ satisfies $E(x)\psi = \|E(x)\|\psi$ only if $P(Z_{\max})\psi = \psi$. The if statement is trivial.

By similar arguments as above, we may show that $\|\mathbb{1}_S - E(x)\| = \|E(\bar{x})\| = p_{\max}(\bar{x}) = 1 - p_{\min}(x)$, and that the unit vector ϕ satisfies $(\mathbb{1}_S - E(x))\phi = \|\mathbb{1}_S - E(x)\|\phi$ if and only if $P(Z_{\min})\phi = \phi$. Since $E(x)$ is non-trivial, then as argued above ψ and ϕ are orthogonal, and perfectly distinguishable by a G measurement. By Proposition M.2 it holds that $\mathcal{I}_{\mathcal{X}}(P_\psi)$ and $\mathcal{I}_{\mathcal{X}}(P_\phi)$ are also perfectly distinguishable by a G measurement, that is, $F(\mathcal{I}_{\mathcal{X}}(P_\psi), \mathcal{I}_{\mathcal{X}}(P_\phi)) = 0$.

□

M.4 Measurements of the first kind, distinguishability, and the Wigner-Araki-Yanase theorem

We shall now use the results in the preceding section to obtain quantitative bounds for first-kind measurements in the presence of a conservation law, that complement our generalisation of the WAY theorem given in Theorem 3.3. To this end, let us first provide a generalisation of Theorem 2 in Ref. [27], which we shall use in the sequel:

Lemma M.5. *Let $\mathcal{M} := (\mathcal{H}_A, \xi, \mathcal{E}, Z)$ be a measurement scheme for an E -instrument \mathcal{I} acting in \mathcal{H}_S , and assume that \mathcal{E} conserves an additive quantity $N = N_S \otimes \mathbb{1}_A + \mathbb{1}_S \otimes N_A$ on average, where $N_S \in \mathcal{L}_s(\mathcal{H}_S)$ and $N_A \in \mathcal{L}_s(\mathcal{H}_A)$. For any pair of orthogonal unit vectors $\psi, \phi \in \mathcal{H}_S$, the following will hold:*

$$|\langle \psi | N_S \phi \rangle| \leq \|N_A\| F(\mathcal{I}_{\mathcal{X}}(P_\psi), \mathcal{I}_{\mathcal{X}}(P_\phi)) + \|N_S\| F(\Lambda(P_\psi), \Lambda(P_\phi)), \quad (59)$$

where Λ is the conjugate channel to $\mathcal{I}_{\mathcal{X}}$ defined in Eq. (5), and $F(\rho, \sigma)$ is the fidelity between states ρ and σ .

Proof. Let us consider the augmented Hilbert space $\mathcal{H}_A \otimes \mathcal{K}$ so that $\xi \in \mathcal{S}(\mathcal{H}_A)$ admits the purification $\xi = \text{tr}_{\mathcal{K}}[P_\varphi]$, with the unit vector $\varphi \in \mathcal{H}_A \otimes \mathcal{K}$. Moreover, if \mathcal{K} is sufficiently large, then by Stinespring's dilation theorem the channel \mathcal{E}^* can be expressed as $\mathcal{E}^*(A) = V^*(A \otimes \mathbb{1}_{\mathcal{K}})V$ for all $A \in \mathcal{L}(\mathcal{H}_S \otimes \mathcal{H}_A)$, where $V : \mathcal{H}_S \otimes \mathcal{H}_A \rightarrow \mathcal{H}_S \otimes \mathcal{H}_A \otimes \mathcal{K}$ is an isometry. By additivity of N , and orthogonality of ψ, ϕ , we have

$$\langle \psi \otimes \varphi | N \phi \otimes \varphi \rangle = \langle \psi | N_S \phi \rangle.$$

On the other hand, average conservation of N by \mathcal{E} implies that

$$N = \mathcal{E}^*(N) = V^*(N_S \otimes \mathbb{1}_A \otimes \mathbb{1}_{\mathcal{K}})V + V^*(\mathbb{1}_S \otimes N_A \otimes \mathbb{1}_{\mathcal{K}})V.$$

We therefore have

$$|\langle \psi | N_S \phi \rangle| \leq |\langle \psi \otimes \varphi | V^*(N_S \otimes \mathbb{1}_A \otimes \mathbb{1}_{\mathcal{K}})V \phi \otimes \varphi \rangle| + |\langle \psi \otimes \varphi | V^*(\mathbb{1}_S \otimes N_A \otimes \mathbb{1}_{\mathcal{K}})V \phi \otimes \varphi \rangle|. \quad (60)$$

For any observable $G := \{G(z)\}$ acting in \mathcal{H}_A , we may write

$$\begin{aligned} & |\langle \psi \otimes \varphi | V^*(N_S \otimes \mathbb{1}_A \otimes \mathbb{1}_{\mathcal{K}})V \phi \otimes \varphi \rangle| \\ &= \left| \sum_z \langle \psi \otimes \varphi | V^*(N_S \otimes G(z) \otimes \mathbb{1}_{\mathcal{K}})V \phi \otimes \varphi \rangle \right| \\ &\leq \|N_S\| \sum_z |\langle \psi \otimes \varphi | V^*(\mathbb{1}_S \otimes G(z) \otimes \mathbb{1}_{\mathcal{K}})V \phi \otimes \varphi \rangle| \\ &\leq \|N_S\| \sum_z |\langle \psi \otimes \varphi | V^*(\mathbb{1}_S \otimes G(z) \otimes \mathbb{1}_{\mathcal{K}})V \psi \otimes \varphi \rangle|^{\frac{1}{2}} |\langle \phi \otimes \varphi | V^*(\mathbb{1}_S \otimes G(z) \otimes \mathbb{1}_{\mathcal{K}})V \phi \otimes \varphi \rangle|^{\frac{1}{2}} \\ &= \|N_S\| \sum_z \text{tr}[\mathbb{1}_S \otimes G(z) \mathcal{E}(P_\psi \otimes \xi)]^{\frac{1}{2}} \text{tr}[\mathbb{1}_S \otimes G(z) \mathcal{E}(P_\phi \otimes \xi)]^{\frac{1}{2}} \\ &= \|N_S\| \sum_z \text{tr}[G(z) \Lambda(P_\psi)]^{\frac{1}{2}} \text{tr}[G(z) \Lambda(P_\phi)]^{\frac{1}{2}}. \end{aligned}$$

In the third line we have used the Cauchy-Schwarz inequality, in the fourth line we used Stinespring's dilation theorem together with the fact that φ is a purification of ξ , and in the final line we use the definitions of the partial trace and the conjugate channel Λ . Now, note that the fidelity satisfies $F(\rho, \sigma) = \min_G \sum_z \text{tr}[G(z)\rho]^{\frac{1}{2}} \text{tr}[G(z)\sigma]^{\frac{1}{2}}$ [95, 96]. Therefore, choosing G so as to obtain the fidelity, we have

$$|\langle \psi \otimes \varphi | V^*(N_S \otimes \mathbb{1}_A \otimes \mathbb{1}_{\mathcal{K}})V \phi \otimes \varphi \rangle| \leq \|N_S\| F(\Lambda(P_\psi), \Lambda(P_\phi)).$$

Using similar steps, we may also write

$$|\langle \psi \otimes \varphi | V^*(\mathbb{1}_S \otimes N_A \otimes \mathbb{1}_{\mathcal{K}})V \phi \otimes \varphi \rangle| \leq \|N_A\| F(\mathcal{I}_{\mathcal{X}}(P_\psi), \mathcal{I}_{\mathcal{X}}(P_\phi)).$$

By Eq. (60), we thus obtain Eq. (59). □

We are now ready to prove our main result in this section:

Theorem M.3. *Consider a measurement scheme $\mathcal{M} := (\mathcal{H}_A, \xi, \mathcal{E}, \mathcal{Z})$ for a nontrivial observable \mathbf{E} with the instrument \mathcal{I} acting in \mathcal{H}_S . Assume that \mathcal{I} is a measurement of the first kind, and that \mathcal{E} conserves an additive quantity $N = N_S \otimes \mathbb{1}_A + \mathbb{1}_S \otimes N_A$ on average, where $N_S \in \mathcal{L}_s(\mathcal{H}_S)$ and $N_A \in \mathcal{L}_s(\mathcal{H}_A)$. For each outcome x associated with a non-trivial effect $\mathbf{E}(x)$, let $\mathcal{K}_{\max}(x)$ and $\mathcal{K}_{\min}(x)$ be subspaces of \mathcal{H}_S defined by*

$$\mathcal{K}_{\max}(x) := \{\psi \in \mathcal{H}_S : \mathbf{E}(x)\psi = \|\mathbf{E}(x)\|\psi\}, \quad \mathcal{K}_{\min}(x) := \{\phi \in \mathcal{H}_S : (\mathbb{1}_S - \mathbf{E}(x))\phi = \|\mathbb{1}_S - \mathbf{E}(x)\|\phi\}.$$

$\mathcal{K}_{\max}(x)$ and $\mathcal{K}_{\min}(x)$ are orthogonal, and for all unit vectors $\psi \in \mathcal{K}_{\max}(x)$ and $\phi \in \mathcal{K}_{\min}(x)$, it holds that

$$|\langle \psi | N_S \phi \rangle| \leq \|N_S\| \left(\|\mathbf{E}(x)\|^{\frac{1}{2}} (1 - \|\mathbb{1}_S - \mathbf{E}(x)\|)^{\frac{1}{2}} + (1 - \|\mathbf{E}(x)\|)^{\frac{1}{2}} \|\mathbb{1}_S - \mathbf{E}(x)\|^{\frac{1}{2}} \right). \quad (61)$$

Proof. For each outcome x associated with a non-trivial effect $\mathbf{E}(x)$, we may coarse-grain \mathbf{E} into a binary observable $\{\mathbf{E}(x), \mathbf{E}(\bar{x}) := \mathbb{1}_S - \mathbf{E}(x)\}$. By Corollary M.2, it holds that for any unit vectors $\psi \in \mathcal{K}_{\max}(x)$ and $\phi \in \mathcal{K}_{\min}(x)$, ψ and ϕ are orthogonal—implying that $\mathcal{K}_{\max}(x)$ and $\mathcal{K}_{\min}(x)$ are orthogonal subspaces—and $F(\mathcal{I}_{\mathcal{X}}(P_\psi), \mathcal{I}_{\mathcal{X}}(P_\phi)) = 0$. As such, given the average conservation of N by the interaction channel \mathcal{E} , Lemma M.5 implies that the following inequality must hold:

$$\begin{aligned} |\langle \psi | N_S \phi \rangle| &\leq \|N_S\| F(\Lambda(P_\psi), \Lambda(P_\phi)) \\ &\leq \|N_S\| \sum_{a=x, \bar{x}} \text{tr}[\mathbf{Z}(a)\Lambda(P_\psi)]^{\frac{1}{2}} \text{tr}[\mathbf{Z}(a)\Lambda(P_\phi)]^{\frac{1}{2}} \\ &= \|N_S\| \sum_{a=x, \bar{x}} \langle \psi | \mathbf{E}(a)\psi \rangle^{\frac{1}{2}} \langle \phi | \mathbf{E}(a)\phi \rangle^{\frac{1}{2}} \\ &= \|N_S\| \left(\|\mathbf{E}(x)\|^{\frac{1}{2}} (1 - \|\mathbb{1}_S - \mathbf{E}(x)\|)^{\frac{1}{2}} + (1 - \|\mathbf{E}(x)\|)^{\frac{1}{2}} \|\mathbb{1}_S - \mathbf{E}(x)\|^{\frac{1}{2}} \right). \end{aligned}$$

The second line uses the fact that for any states ρ, σ , it holds that $F(\rho, \sigma) \leq \sum_a \text{tr}[\mathbf{F}(a)\rho]^{\frac{1}{2}} \text{tr}[\mathbf{F}(a)\sigma]^{\frac{1}{2}}$ for any observable \mathbf{F} [95, 96]. The third line uses the fact that Λ is the conjugate channel to $\mathcal{I}_{\mathcal{X}}$ defined in Eq. (5), and so it holds that $\text{tr}[\mathbf{Z}(a)\Lambda(\rho)] = \text{tr}[\mathbf{E}(a)\rho]$ for all ρ and $a = x, \bar{x}$. To see how the final line is obtained, note that we have $\langle \psi | \mathbf{E}(x)\psi \rangle = \|\mathbf{E}(x)\|$ and $\langle \phi | (\mathbb{1}_S - \mathbf{E}(x))\phi \rangle = \|\mathbb{1}_S - \mathbf{E}(x)\|$ by construction. For the first term, i.e., $a = x$, we obtain $\langle \psi | \mathbf{E}(a)\psi \rangle = \|\mathbf{E}(x)\|$ and $\langle \phi | \mathbf{E}(a)\phi \rangle = \langle \phi | (\mathbb{1}_S - (\mathbb{1}_S - \mathbf{E}(x)))\phi \rangle = 1 - \langle \phi | (\mathbb{1}_S - \mathbf{E}(x))\phi \rangle = 1 - \|\mathbb{1}_S - \mathbf{E}(x)\|$. The second term for $a = \bar{x}$ is obtained in a similar manner. \square

Let us note that if \mathbf{E} commutes with N_S , then Theorem M.3 imposes no restrictions on first-kindness. To see this, let us note that for any (possibly trivial) effect $\mathbf{E}(x)$, and for any unit vectors ψ, ϕ satisfying $\mathbf{E}(x)\psi = \|\mathbf{E}(x)\|\psi$ and $(\mathbb{1}_S - \mathbf{E}(x))\phi = \|\mathbb{1}_S - \mathbf{E}(x)\|\phi$, it holds that $\langle \psi | \mathbf{E}(x)N_S\phi \rangle = \|\mathbf{E}(x)\|\langle \psi | N_S\phi \rangle$ and $\langle \psi | N_S\mathbf{E}(x)\phi \rangle = (1 - \|\mathbb{1}_S - \mathbf{E}(x)\|)\langle \psi | N_S\phi \rangle$. It follows that if $[\mathbf{E}(x), N_S] = \mathbf{0}$ then either (i) $\langle \psi | N_S\phi \rangle = 0$, or (ii) $\|\mathbf{E}(x)\| + \|\mathbb{1}_S - \mathbf{E}(x)\| = 1$. Condition (i) implies that the lower bound of Eq. (61) vanishes, and so no constraint is imposed. On the other hand, condition (ii) implies that $\|\mathbf{E}(x)\| + \|\mathbb{1}_S - \mathbf{E}(x)\| = 1 + (p_{\max}(x) - p_{\min}(x)) = 1$, where we recall from Corollary M.2 that $p_{\max}(x) = \|\mathbf{E}(x)\|$ and $p_{\min}(x) = 1 - \|\mathbb{1}_S - \mathbf{E}(x)\|$ are the largest and smallest values from the set $\{p(x|z)\}$ given by Theorem M.2. Such equality is satisfied if and only if $p_{\max}(x) = p_{\min}(x) = \lambda$, in which case by Eq. (58) it follows that $\mathbf{E}(x) = \lambda\mathbb{1}_S$ is a trivial effect. But Eq. (61) applies only to non-trivial effects, and so no constraints are imposed in such a case.

Let us now show that the first-kindness statement of Theorem 3.3 may be recovered directly from Eq. (61).

Corollary M.3. *Consider a measurement scheme $\mathcal{M} := (\mathcal{H}_A, \xi, \mathcal{E}, \mathcal{Z})$ for an \mathbf{E} -instrument \mathcal{I} acting in \mathcal{H}_S , and assume that \mathcal{E} conserves an additive quantity $N = N_S \otimes \mathbb{1}_A + \mathbb{1}_S \otimes N_A$ on average, where $N_S \in \mathcal{L}_s(\mathcal{H}_S)$ and $N_A \in \mathcal{L}_s(\mathcal{H}_A)$. If \mathcal{I} is a measurement of the first kind, then for all effects $\mathbf{E}(x)$ that have both eigenvalue 1 and 0,*

$$\mathbf{P}[\mathbf{E}(x), N_S]\mathbf{P} = [\mathbf{E}(x), \mathbf{P}N_S\mathbf{P}] = \mathbf{0},$$

where $\mathbf{P} = \mathbf{P}_0(x) + \mathbf{P}_1(x)$, with $\mathbf{P}_0(x)$ and $\mathbf{P}_1(x)$ orthogonal projections onto the eigenvalue-0 and eigenvalue-1 eigenspaces of $\mathbf{E}(x)$, respectively.

Proof. It is trivial to verify that $\text{supp}(\mathbf{P}_1(x)) \equiv \mathcal{K}_{\max}(x)$ and $\text{supp}(\mathbf{P}_0(x)) \equiv \mathcal{K}_{\min}(x)$ as defined in Theorem M.3. Any $\varphi \in \text{supp}(\mathbf{P})$ may be written as $\varphi = \alpha\psi + \beta\phi$, where $\psi \in \mathcal{K}_{\max}(x)$ and $\phi \in \mathcal{K}_{\min}(x)$ are unit vectors and $\alpha, \beta \in \mathbb{C}$. Since $\|\mathbf{E}(x)\| = \|\mathbb{1}_S - \mathbf{E}(x)\| = 1$ for any effect that has both eigenvalue 1 and 0, then by Theorem M.3

it follows that $\langle \varphi | P_1(x) N_S P_0(x) \varphi \rangle = \alpha^* \beta \langle \psi | N_S \phi \rangle = 0$ for all $\varphi \in \text{supp}(P)$. It follows that $P_1(x) N_S P_0(x) = P_1(x) N_S (P - P_1(x)) = \mathbf{0}$, which implies that $P_1(x) N_S P = P_1(x) N_S P_1(x)$. Since the right hand side is self-adjoint, and $P_1(x) = P_1(x) P = P P_1(x)$, it follows that $[P_1(x), P N_S P] = \mathbf{0}$. But since $P_1(x) = E(x) P = P E(x)$, we have $[E(x), P N_S P] = P [E(x), N_S] P = \mathbf{0}$. This completes the proof. \square

References

- [1] P. Busch, G. Cassinelli, and P. J. Lahti, *Found. Phys.* **20**, 757 (1990).
- [2] M. Ozawa, *Phys. Rev. A* **67**, 042105 (2003).
- [3] P. Busch, in *Quantum Reality, Relativ. Causality, Closing Epistemic Circ.* (Springer, Dordrecht, 2009) pp. 229–256.
- [4] T. Heinosaari and M. M. Wolf, *J. Math. Phys.* **51**, 092201 (2010).
- [5] M. Tsang and C. M. Caves, *Phys. Rev. Lett.* **105**, 123601 (2010).
- [6] M. Tsang and C. M. Caves, *Phys. Rev. X* **2**, 1 (2012).
- [7] L. A. Rozema, A. Darabi, D. H. Mahler, A. Hayat, Y. Soudagar, and A. M. Steinberg, *Phys. Rev. Lett.* **109**, 100404 (2012).
- [8] J. P. Groen, D. Ristè, L. Tornberg, J. Cramer, P. C. de Groot, T. Picot, G. Johansson, and L. DiCarlo, *Phys. Rev. Lett.* **111**, 090506 (2013).
- [9] M. Hatridge, S. Shankar, M. Mirrahimi, F. Schackert, K. Geerlings, T. Brecht, K. M. Sliwa, B. Abdo, L. Frunzio, S. M. Girvin, R. J. Schoelkopf, and M. H. Devoret, *Science* (80-.). **339**, 178 (2013).
- [10] P. Busch, P. Lahti, and R. F. Werner, *Phys. Rev. Lett.* **111**, 160405 (2013).
- [11] P. Busch, P. Lahti, and R. F. Werner, *Rev. Mod. Phys.* **86**, 1261 (2014).
- [12] F. Kaneda, S.-Y. Baek, M. Ozawa, and K. Edamatsu, *Phys. Rev. Lett.* **112**, 020402 (2014).
- [13] M. S. Blok, C. Bonato, M. L. Markham, D. J. Twitchen, V. V. Dobrovitski, and R. Hanson, *Nat. Phys.* **10**, 189 (2014).
- [14] T. Shitara, Y. Kuramochi, and M. Ueda, *Phys. Rev. A* **93**, 032134 (2016).
- [15] C. B. Møller, R. A. Thomas, G. Vasilakis, E. Zeuthen, Y. Tsaturyan, M. Balabas, K. Jensen, A. Schliesser, K. Hammerer, and E. S. Polzik, *Nature* **547**, 191 (2017).
- [16] I. Hamamura and T. Miyadera, *J. Math. Phys.* **60**, 082103 (2019).
- [17] C. Carmeli, T. Heinosaari, T. Miyadera, and A. Toigo, *Found. Phys.* **49**, 492 (2019).
- [18] K.-D. Wu, E. Bäumer, J.-F. Tang, K. V. Hovhannisyanyan, M. Perarnau-Llobet, G.-Y. Xiang, C.-F. Li, and G.-C. Guo, *Phys. Rev. Lett.* **125**, 210401 (2020).
- [19] G. M. D’Ariano, P. Perinotti, and A. Tosini, *Quantum* **4**, 363 (2020).
- [20] A. C. Ipsen, *Found. Phys.* **52**, 20 (2022).
- [21] T. Heinosaari, T. Miyadera, and M. Ziman, *J. Phys. A Math. Theor.* **49**, 123001 (2016).
- [22] O. Gühne, E. Haapasalo, T. Kraft, J.-P. Pellonpää, and R. Uola, *Rev. Mod. Phys.* **95**, 011003 (2023).
- [23] E. P. Wigner, *Zeitschrift für Phys. A Hadron. Nucl.* **133**, 101 (1952).
- [24] P. Busch, (2010), arXiv:1012.4372 .
- [25] H. Araki and M. M. Yanase, *Phys. Rev.* **120**, 622 (1960).
- [26] L. Loveridge and P. Busch, *Eur. Phys. J. D* **62**, 297 (2011).
- [27] T. Miyadera and H. Imai, *Phys. Rev. A* **74**, 024101 (2006).
- [28] G. Kimura, B. Meister, and M. Ozawa, *Phys. Rev. A* **78**, 032106 (2008).
- [29] P. Busch and L. Loveridge, *Phys. Rev. Lett.* **106**, 110406 (2011).
- [30] P. Busch and L. D. Loveridge, in *Symmetries Groups Contemp. Phys.* (WORLD SCIENTIFIC, 2013) pp. 587–592.
- [31] A. Luczak, *Open Syst. Inf. Dyn.* **23**, 1 (2016).
- [32] M. Tukiainen, *Phys. Rev. A* **95**, 012127 (2017).
- [33] H. Tajima and H. Nagaoka, (2019), arXiv:1909.02904 .
- [34] S. Soitan, M. Frączak, W. Belzig, and A. Bednorz, *Phys. Rev. Res.* **3**, 013247 (2021).
- [35] M. Ozawa, *Phys. Rev. Lett.* **89**, 3 (2002).
- [36] T. Karasawa and M. Ozawa, *Phys. Rev. A* **75**, 032324 (2007).
- [37] T. Karasawa, J. Gea-Banacloche, and M. Ozawa, *J. Phys. A Math. Theor.* **42**, 225303 (2009).
- [38] M. Ahmadi, D. Jennings, and T. Rudolph, *New J. Phys.* **15**, 013057 (2013).
- [39] J. Åberg, *Phys. Rev. Lett.* **113**, 150402 (2014).
- [40] H. Tajima, N. Shiraishi, and K. Saito, *Phys. Rev. Res.* **2**, 043374 (2020).
- [41] L. Loveridge, T. Miyadera, and P. Busch, *Found. Phys.* **48**, 135 (2018).
- [42] L. Loveridge, *J. Phys. Conf. Ser.* **1638**, 012009 (2020).
- [43] N. Gisin and E. Zambrini Cruzeiro, *Ann. Phys.* **530**, 1700388 (2018).

- [44] M. Navascués and S. Popescu, *Phys. Rev. Lett.* **112**, 140502 (2014).
- [45] M. H. Mohammady and J. Anders, *New J. Phys.* **19**, 113026 (2017).
- [46] M. H. Mohammady and A. Romito, *Quantum* **3**, 175 (2019).
- [47] G. Chiribella, Y. Yang, and R. Renner, *Phys. Rev. X* **11**, 021014 (2021).
- [48] M. H. Mohammady, *Phys. Rev. A* **104**, 062202 (2021).
- [49] P. Busch, P. Lahti, J.-P. Pellonpää, and K. Ylinen, *Quantum Measurement*, Theoretical and Mathematical Physics (Springer International Publishing, Cham, 2016).
- [50] P. Busch, M. Grabowski, and P. J. Lahti, *Operational Quantum Physics*, Lecture Notes in Physics Monographs, Vol. 31 (Springer Berlin Heidelberg, Berlin, Heidelberg, 1995).
- [51] P. Busch, P. J. Lahti, and Peter Mittelstaedt, *The Quantum Theory of Measurement*, Lecture Notes in Physics Monographs, Vol. 2 (Springer Berlin Heidelberg, Berlin, Heidelberg, 1996).
- [52] T. Heinosaari and M. Ziman, *The Mathematical language of Quantum Theory* (Cambridge University Press, Cambridge, 2011).
- [53] B. Janssens, *Lett. Math. Phys.* **107**, 1557 (2017).
- [54] O. Bratteli and D. W. Robinson, *Operator Algebras and Quantum Statistical Mechanics 1* (Springer Berlin Heidelberg, Berlin, Heidelberg, 1987).
- [55] O. Bratteli, P. E. T. Jorgensen, A. Kishimoto, and R. F. Werner, *J. Oper. Theory* **43**, 97 (2000).
- [56] E. B. Davies and J. T. Lewis, *Commun. Math. Phys.* **17**, 239 (1970).
- [57] M. Ozawa, *Phys. Rev. A* **62**, 062101 (2000).
- [58] M. Ozawa, *Phys. Rev. A* **63**, 032109 (2001).
- [59] J.-P. Pellonpää, *J. Phys. A Math. Theor.* **46**, 025302 (2013).
- [60] J.-P. Pellonpää, *J. Phys. A Math. Theor.* **46**, 025303 (2013).
- [61] G. Lüders, *Ann. Phys.* **518**, 663 (2006).
- [62] M. Ozawa, *J. Math. Phys.* **25**, 79 (1984).
- [63] P. Busch and J. Singh, *Phys. Lett. A* **249**, 10 (1998).
- [64] P. Busch, M. Grabowski, and P. J. Lahti, *Found. Phys.* **25**, 1239 (1995).
- [65] P. J. Lahti, P. Busch, and P. Mittelstaedt, *J. Math. Phys.* **32**, 2770 (1991).
- [66] M. M. Yanase, *Phys. Rev.* **123**, 666 (1961).
- [67] M. Ozawa, *Phys. Rev. Lett.* **88**, 050402 (2002).
- [68] I. Marvian and R. W. Spekkens, *Nat. Commun.* **5**, 3821 (2014).
- [69] C. Cîrstoiu, K. Korzekwa, and D. Jennings, *Phys. Rev. X* **10**, 041035 (2020).
- [70] D. Petz and C. Ghinea, *Quantum Probab. Relat. Top.* (World Scientific, Singapore, 2011) pp. 261–281.
- [71] A. Streltsov, G. Adesso, and M. B. Plenio, *Rev. Mod. Phys.* **89**, 041003 (2017).
- [72] R. Takagi, *Sci. Rep.* **9**, 14562 (2019).
- [73] I. Marvian, *Phys. Rev. Lett.* **129**, 190502 (2022).
- [74] G. Tóth and D. Petz, *Phys. Rev. A* **87**, 032324 (2013).
- [75] S. Yu, (2013), arXiv:1302.5311 .
- [76] L. Weihua and W. Junde, *J. Phys. A Math. Theor.* **43**, 395206 (2010).
- [77] B. Prunaru, *J. Phys. A Math. Theor.* **44**, 185203 (2011).
- [78] A. Arias, A. Gheondea, and S. Gudder, *J. Math. Phys.* **43**, 5872 (2002).
- [79] L. Weihua and W. Junde, *J. Math. Phys.* **50**, 103531 (2009).
- [80] G. M. D’Ariano, P. Perinotti, and M. Sedlák, *J. Math. Phys.* **52**, 082202 (2011).
- [81] M. H. Mohammady, *Phys. Rev. A* **103**, 042214 (2021).
- [82] V. Pata, *Fixed Point Theorems and Applications*, UNITEXT, Vol. 116 (Springer International Publishing, Cham, 2019).
- [83] G. Pisier, *Introduction to Operator Space Theory* (Cambridge University Press, 2003).
- [84] Y. Kuramochi and H. Tajima, (2022), arXiv:2208.13494 .
- [85] R. V. Kadison, *Ann. Math.* **56**, 494 (1952).
- [86] M.-D. Choi, *Illinois J. Math.* **18**, 565 (1974).
- [87] W. F. Stinespring, *Proc. Am. Math. Soc.* **6**, 211 (1955).
- [88] T. Miyadera and H. Imai, *Phys. Rev. A* **78**, 052119 (2008).
- [89] T. Miyadera, L. Loveridge, and P. Busch, *J. Phys. A Math. Theor.* **49**, 185301 (2016).
- [90] K. Kraus, *States, Effects, and Operations Fundamental Notions of Quantum Theory*, edited by K. Kraus, A. Böhm, J. D. Dollard, and W. H. Wootters, Lecture Notes in Physics, Vol. 190 (Springer Berlin Heidelberg, Berlin, Heidelberg, 1983).
- [91] P. Lahti, *Int. J. Theor. Phys.* **42**, 893 (2003).
- [92] J.-P. Pellonpää, *J. Phys. A Math. Theor.* **47**, 052002 (2014).
- [93] S. Luo and Q. Zhang, *Theor. Math. Phys.* **151**, 529 (2007).

- [94] G. M. D'Ariano, P. L. Presti, and P. Perinotti, *J. Phys. A: Math. Gen.* **38**, 5979 (2005).
- [95] C. A. Fuchs and C. M. Caves, *Open Syst. Inf. Dyn.* **3**, 345 (1995).
- [96] H. Barnum, C. M. Caves, C. A. Fuchs, R. Jozsa, and B. Schumacher, *Phys. Rev. Lett.* **76**, 2818 (1996).