Imperial College London
Centre for Doctoral Training in Controlled Quantum Dynamics
Department of Physics

Resource theory of asymmetry

Kamil Korzekwa

September 20, 2013

Supervised by Dr David Jennings and Prof. Terry Rudolph

Report submitted in partial fulfilment of the requirements
of the degree of MRes of Imperial College London

Abstract

This thesis deals with the subject of the error-disturbance trade-off relation in quantum measurements using a new quantum information approach based on the resource theory of asymmetry. The goal is to fully understand and quantify the fundamental restrictions imposed on the information about a quantum state available through measurements. The connection between the symmetric component of a quantum stat, with respect to the U(1) symmetry generated by a given observable, and the measurement process of this observable is presented and the asymmetry resource of a state is linked with the state-dependent error-disturbance uncertainty relation. The requirements for physically meaningful definitions of a state-dependent measurement error and disturbance are proposed and a detailed discussion of the current approaches to the subject of sequential measurements is given. It is pointed out that these approaches fail to correctly describe the trade-off between the information gained through the measurement of one observable and the disturbance of the measurement of the other. Then the limitation on quantum measurements induced by conservation laws, the famous Wigner-Araki-Yanase (WAY) theorem, is analyzed, both from a historical and contemporary research perspective. An asymmetry resource formulation of this problem is also presented, where the restriction of symmetric evolution limits the insight into the asymmetric part of the quantum state. Next, the current bounds for both the error-disturbance relation and the noise in the WAY scenario measurement are improved with the help of the generalized Heisenberg uncertainty relations for mixed states. Finally these generalized relations together with the entropic uncertainty relations are used to bound the asymmetric properties of a quantum state with respect to two non-commuting observables.

Acknowledgements

I would like to thank my supervisors, Dr David Jennings and Prof. Terry Rudolph, for their friendly guidance, support and wise advices. I would also like to thank all my friends from the Cohort 4 of the CQD CDT for being such great companions during the last year and especially Matteo Lostaglio for many fruitful discussions. Last, but not least, I would like to express my gratitude and thanks to the staff and directors of the CQD CDT for creating a great atmosphere and providing not only the financial support, but everything that is needed to work comfortably.

Contents

Introduction

Despite almost a century of research on quantum theory, one of its fundamental building blocks, the quantum measurement process, is still not well understood. One of the well-known results in this field is the famous Heisenberg uncertainty relation. Its best known modern formulation (also known as the Heisenberg-Robertson uncertainty relation [1]) concerns the outcome statistics of two independent measurements of non-commuting observables performed on an ensemble of identically prepared quantum states. It states that the product of variances of these two outcome statistics is lower-bounded by the mean value of the commutator of measured observables in the given quantum state. Although this formulation says nothing about the effect of one measurement on the outcome statistics of the other, it is usually misinterpreted in the spirit of the original Heisenberg microscope thought experiment, i.e. that the bigger the precision of the measurement of one observable, the bigger the disturbance of the measurement of the other one (not commuting with the first one). Indeed, the formulation of uncertainty relation in terms of precision and disturbance of sequential measurements was the original Heisenberg's concept and is called the noise-disturbance uncertainty relation. Interestingly it was not until recently that the problem of sequential measurements was addressed with mathematically rigorous approach [2]. However, the problem is not yet fully understood, mainly due to the discussion over the proper definitions of noise and disturbance for quantum measurements.

Another interesting, however not so widely known, result limiting the measurement precision is the Wigner-Araki-Yanase (WAY) theorem [3, 4]. It states that there exists no exact measurement of a quantum operator that does not commute with a conserved quantity operator. Since it is well known that the presence of a conservation law results from the underlying continous symmetry of the system evolution (via the Noether's theorem [5]), one can restate the WAY theorem: measurement of an operator which is not invariant under the symmetry operation of the system evolution cannot be exact. Although exact measurements are not allowed, the approximate measurement processes exist. Recently, Ozawa found the lower bound for the error of such measurements [6] by using a variant of his generalized noise-disturbance relation [2]. The origin of the bound, however, and especially its dependence on the uncertainty (variance) of the conserved quantity in the initial states of the measured system and measuring apparatus is still not well understood.

Both the noise-disturbance uncertainty relation (NDUR) and the WAY theorem may be seen as restrictions on the information about the quantum state available to the observer. The NDUR states that obtaining the information on the quantum state distribution over eigenstates of one observable (by measuring it) disturbs the distribution over the eigenstates of the other observable, so that the full information abut the initial quantum state cannot be recovered. The WAY theorem on the other hand limits the insight into the asymmetric (with respect to U(1) symmetry generated by the conserved quantity) part of the quantum state.

The aim of this project is to create a new, information-theoretic description of the quantum measurement process that would allow to fully understand and quantify the fundamental restrictions imposed on the information about the quantum state available through measurements. The resource theory of asymmetry [7, 8] is a well-suited tool for this task. As will be shown the measurement of a quantum observable A is equivalent to the projection onto the symmetric subspace generated by A (the so-called $\mathcal{G}$-twirling). Therefore the investigation of relations between symmetric subspaces for two non-commuting observables can shed light on the noise-disturbance relation for sequential measurements. Also, in the case of the WAY scenario measurements, due to the formulation in terms of symmetric evolution, the resource theory of asymmetry seems to be a natural language to work with.

The thesis consists of four chapters. In the first one, the basic concepts and ideas concerning the resource theory of asymmetry are presented, as well as its connection to the measurement process. Next, in Chapter 2, the problem of sequential measurements is discussed. The arguments for using the definitions of noise and disturbance based solely on the outcome statistics of measurements are first presented and then the current approaches [2, 9] to quantify and bound error and disturbance are analyzed and compared with the proposed ones. Chapter 3 addresses the problem of quantum measurements in the presence of a conservation law (the WAY scenario measurements). The historical overview of this subject is given together with the recent approach of Ozawa [6] and the formulation of the problem in terms of the asymmetry resources [10, 11]. Finally, Chapter 4 contains applications of different results from the general field of uncertainty relations to problems presented in the first three chapters.

Chapter 1

Asymmetry resources

All resource theories are formulated in terms of free states and free transformations. Free transformations are assumed to be the only allowed transformations of the theory and they map the set of free states into itself. Therefore every resource theory can be formulated by the restriction for the allowed transformations and the resulting set of free states. For example, in the resource formulation of entanglement theory free transformations are restricted to the local operations and classical communication (LOCC) and the set of free states is composed of separable (unentangled) states. Every state that is not free is considered to be a resource. The resource theory studies the manipulation of resource states under the free transformations. This includes e.g. the possibility of interconversion between states, partial ordering based on some measure of resource or studying the advantages of possesing a resource state in quantum communication or information processing.

In the resource theory of asymmetry [8] (also known as the theory of quantum reference frames) asymmetric states are treated as resources under the limitation of the symmetric evolution, i.e. being restricted to symmetric evolution one cannot create asymmetric state, but having one allows for otherwise forbidden operations (as asymmetric states, unlike the symmetric ones, are affected by symmetric evolution). In this chapter we will first present the basic definitions and theorems of the general theory and then focus on some specific results concerning the U(1) asymmetry resource theory. We will also connect the formalism of the U(1) asymmetry with the quantum measurement process, specifically with the information gain-disturbance trade-off relation.

1.1 Basic definitions and theorems

The set of free transformations is defined by transformations that are symmetric with respect to some symmetry group G. Formally it means that the set is restricted to quantum operations $\mathcal{E}$ that are covariant under the unitary representation of the group, $\{U(g), \forall g \in G\}$, i.e.

$$\forall g \in G : \mathcal{E}[U(g)(\cdot)U^\dagger(g)] = U(g)\mathcal{E}(\cdot)U^\dagger(g), \tag{1.1}$$

which can be equivalently written in terms of vanishing commutators,

$$\forall g \in G : [\mathcal{E}, \mathcal{U}(g)] = 0, \tag{1.2}$$

where $\mathcal{U}(g)(\cdot) = U(g)(\cdot)U^\dagger(g)$. Similarly, the set of free states is composed of states ρ that are invariant under the unitary representation of the group G, i.e. that are symmetric with respect to symmetry group G,

$$\forall g \in G : U(g)\rho U^\dagger(g) = \rho, \tag{1.3}$$

or equivalently

$$\forall g \in G : [\rho, U(g)] = 0. \tag{1.4}$$

As an example consider the set of unitaries $\exp(i\phi L_z)$, which are the unitary representation of the $U(1)$ group. Since these operators are representing the rotation around the z axis in the position space, the corresponding invariant states must be invariant with respect to these rotations. The set of free states is thus composed of the states diagonal in the basis of the eigenstates of the L_z operator, in particular the subset of pure free states is composed of eigenstates of the L_z operator.

In order to specify the asymmetric properties of a state ρ one can find all its symmetries, the symmetry subgroup of ρ, which is simply defined as the subgroup of G under which ρ is invariant,

$$\mathrm{Sym}_G(\rho) = \{g \in G : U(g)\rho U^\dagger(g) = \rho\}. \tag{1.5}$$

One can clearly see from the above definition and Eq. 1.3 that the symmetry subgroup of a free state is the whole group G. One of the important results concerning the symmetry subgroup of states is that under G-covariant transformation $\mathcal{E}$ all the symmetries of the initial state are preserved [8], i.e.

$$\mathrm{Sym}_G(\rho) \subseteq \mathrm{Sym}_G[\mathcal{E}(\rho)]. \tag{1.6}$$

As a result being restricted to G-covariant operations one cannot create asymmetric states from symmetric ones or transform one kind of asymmetric state into the other (for example invariant only under rotations around x into invariant only under rotations around z).

States that can be reversibly interconverted under the G-covariant operation form the so called G-equivalence class. More formally two states, ρ and σ, belong to the same G-equivalence class, if there exist G-covariant quantum operations $\mathcal{E}$ and $\mathcal{F}$ such that

$$\mathcal{E}(\rho) = \sigma, \tag{1.7a}$$

$$\mathcal{F}(\sigma) = \rho. \tag{1.7b}$$

States belonging to the same G-equivalence class must have all the same symmetries.

The $\mathcal{G}$-twirling operation is defined as the averaging over the group,

$$\mathcal{G}(\rho) = \int dg U(g)\rho U^{\dagger}(g), \tag{1.8a}$$

$$\mathcal{G}(\mathcal{E}(\cdot)) = \int dg U(g)\mathcal{E}(\cdot)U^{\dagger}(g). \tag{1.8b}$$

It can be easily verified that $\mathcal{G}$-twirled state (operation) is a symmetric state (operation). In fact the $\mathcal{G}$-twirling is the projection onto the symmetric subspace of states (operations). As such it leaves symmetric states (operations) invariant, i.e. if ρ ($\mathcal{E}$) is symmetric then $\mathcal{G}(\sigma) = \sigma$ ($\mathcal{G}(\mathcal{E}) = \mathcal{E}$). With the use of $\mathcal{G}$-twirling one can express the natural measure of asymmetry, called the relative entropy of frameness, in a particularly simple form. The relative entropy of frameness is defined as the relative entropy distance of a state ρ to the nearest symmetric state,

$$As_G(\rho) = \min_{\sigma:\ \mathcal{G}(\sigma)=\sigma} \{S(\rho||\sigma)\}. \tag{1.9}$$

One can show [12] that the nearest symmetric state to ρ is just its projection onto the symmetric subspace, so that

$$As_G(\rho) = S(\rho||\mathcal{G}(\rho)) = S(\mathcal{G}(\rho)) - S(\rho). \tag{1.10}$$

In particular case of pure states it simplifies to

$$As_G(|\psi\rangle\langle\psi|) = S(\mathcal{G}(|\psi\rangle\langle\psi|)). \tag{1.11}$$

1.2 U(1) asymmetry resource theory

As will be explained in the next section, the case of $G = U(1)$ is of particular interest for the quantum measurements related considerations, therefore some results concerning it will be presented now. The unitary representation of this group is generated by operator $N = \sum_n n |n\rangle\langle n|$, so that $U(\theta) = \exp(i\theta N)$ and the Hilbert space splits up into charge sectors of N, $\mathcal{H} = \oplus_n \mathcal{H}_n$. The general form of U(1)-invariant unitary is diagonal in the $\{|n\rangle\}$ basis,

$$V_{U(1)} = \sum_n e^{i\phi_n} |n\rangle\langle n|, \tag{1.12}$$

therefore the only action allowed by symmetric unitary is the change of relative phases between the eigenstates of N. The more general U(1)-covariant operations have the following Kraus decomposition [13]

$$\mathcal{E}_{U(1)}(\cdot) = \sum_{k,\alpha} K^{\dagger}_{k,\alpha}(\cdot)K_{k,\alpha}, \tag{1.13a}$$

$$K_{k,\alpha} = S_k\tilde{K}_{k,\alpha}, \tag{1.13b}$$

$$\tilde{K}_{k,\alpha} = \sum_{n} c_n^{(k,\alpha)} |n\rangle\langle n|, \tag{1.13c}$$

$$S_k = \sum_{n=\max\{0,-k\}} |n+k\rangle\langle n|, \tag{1.13d}$$

and allow not only to change the relative phases (the action of $\tilde{K}_{k,\alpha}$), but also to shift the number of excitations (the action of S_k shifts it upwards for $k>0$ and downwards for $k<0$).

Basing on the above result one can find the condition for the pure state $|\psi_1\rangle$ to be convertible into $|\psi_2\rangle$ under symmetric evolution [13]. First note that under symmetric evolution any two states can be brought to the standard form

$$|\psi_1\rangle = \sum_n \sqrt{p_n}\,|n\rangle, \tag{1.14a}$$

$$|\psi_2\rangle = \sum_n \sqrt{q_n}\,|n\rangle. \tag{1.14b}$$

Now the condition for convertibility in terms of the distributions $\mathbf{p}=(p_1,p_2,\dots)$ and $\mathbf{q}=(q_1,q_2,\dots)$ can be written as

$$\mathbf{p}=\sum_k w_k T^{(k)}\mathbf{q}, \tag{1.15}$$

where $0\leq w_k\leq 1$, $\sum_k w_k=1$ and $[T^{(k)}\mathbf{q}]_j=q_{j+k}$.

It is interesting to note that if one limits only to the U(1)-covariant transformations between pure states, then the variance V of N in a given state is an asymmetry monotone. In order to see this, let us first use the concavity of variance (proven in Sec. 4.3.1),

$$V(\mathbf{p})=V\left(\sum w_k T^{(k)}\mathbf{q}\right)\geq\sum_k w_k V\left(T^{(k)}\mathbf{q}\right). \tag{1.16}$$

Now one can show that $V(\mathbf{q})=V(T^{(k)}\mathbf{q})$ in the following way

$$\begin{aligned} V(T^{(k)}\mathbf{q}) &= \sum_n q_n(n+k)^2-\sum_{n,m}q_nq_m(n+k)(m+k)=\sum_n q_n(n^2+2nk+k^2)\\ &\quad -\sum_{n,m}q_nq_m(nm+nk+mk+k^2)=\sum_n q_n n^2-\sum_{n,m}q_nq_m nm=V(\mathbf{q}). \end{aligned}$$

Substituting this result to Eq. 1.16 one gets

$$V(\mathbf{p})\geq\sum_k w_k V(\mathbf{q})=V(\mathbf{q}), \tag{1.17}$$

which means that the variance is monotonically decreasing under U(1)-covariant transformations between pure states. In Chapter 4 it will be shown that the extension of variance as a measure of asymmetry for mixed states is given by the Wigner-Yanase skew information.

The operational meaning of variance V as a measure of asymmetry resource was given by Gour [13]. Defining the maximally asymmetric qubit state $|\phi\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$ one can introduce the measure based on the asymptotic conversion rate into $|\phi\rangle$. Having $M \to \infty$ copies of a given pure state $|\psi\rangle$ one can convert them using U(1)-covariant transformations into $4MV$ copies of $|\phi\rangle$. Therefore the scaled variance of a state is equal to the asymptotic conversion rate of this state into the maximally asymmetric qubit state.

1.3 Connection between projective measurements and U(1) asymmetry resources

The non-selective projective measurement of observable $A = \sum a_n |a_n\rangle \langle a_n|$ on a quantum state ρ is defined as a measurement of A for which the outcome is not recorded. Alternatively one can use the following operational scenario for such measurement: one party, Alice, posesses the quantum state ρ, sends it to another party, Bob, which performs the measurement of A (known to Alice) and sends the state back to her, without informing about the measurement outcome. One can also see the non-selective measurement as the average effect of the measurement on a state: starting from the ensemble of identically prepared states, $\rho^{\otimes N}$, one performs the measurement of A on every copy of ρ, then mixes all the systems together and chooses one at random. Of course all these processes have the same effect and transform the initial state ρ into

$$\rho' = \sum_n |a_n\rangle \langle a_n| \rho |a_n\rangle \langle a_n| = \sum_n \rho_{nn} |a_n\rangle \langle a_n| , \tag{1.18}$$

thus killing all off-diagonal terms of ρ in the $\{|a_n\rangle\}$ basis. As the measurement outcomes cannot be controlled (due to the probabilistic nature of quantum mechanics), when one wants to describe the effect of the projective measurement of A (on the measured state, on the subsequent measurements or on anything else), it is reasonable to use its average effect. For example, if one is looking for the disturbance that the measurement of A induced on the subsequent measurement of B, one should compare the measurements of B on states ρ and ρ'.

Let us now look at the $\mathcal{G}$-twirling operation for $U(1)$ group generated by observable A,

$$\mathcal{G}_A(\cdot) = \frac{1}{2\pi} \int_0^{2\pi} d\phi e^{-i\phi A} (\cdot) e^{i\phi A}, \tag{1.19}$$

and the effect it has on the arbitrary state ρ,

$$\begin{aligned}\mathcal{G}_A(\rho) &= \frac{1}{2\pi}\int_0^{2\pi} d\phi e^{-i\phi A}\left(\sum_{m,n} |a_m\rangle\langle a_m|\rho|a_n\rangle\langle a_n|\right) e^{i\phi A}\\ &= \sum_{m,n}\langle a_m|\rho|a_n\rangle |a_m\rangle\langle a_n| \frac{1}{2\pi}\int_0^{2\pi} d\phi e^{i\phi(n-m)} = \sum_n \langle a_n|\rho|a_n\rangle |a_n\rangle\langle a_n| .\end{aligned}$$

By comparing the above result with Eq. 1.18 one can see that the effect of $\mathcal{G}$-twirling is exactly the same as the effect of the non-selective measurement. Hence the measurement of A can be seen as the projection onto the subspace of symmetric states. This means that if one's insight into the quantum state ρ is restricted only to the measurement outcomes of some observable A then if the state ρ is not symmetric (with respect to symmetry generated by A), one does not have access to the full information about the system. To see this more clearly let us note that the full information about a state ρ living in a d-dimensional Hilbert space requires the specification of (d^2-1) real parameters, whereas the symmetric component is fully specified by the outcome probability distribution, so by $(d-1)$ real parameters. This is exactly the complementarity principle: measurement of a given observable A of a system in the state ρ gives the observer only the insight into one of the classical contexts of ρ, the one connected with definite values of A.

If one wants to get the full information on an unkown state ρ, a state tomography must be performed, i.e. one has to perform measurements of different non-commuting observables to obtain (d^2-1) real parameters. However in order to do this one needs an access to the ensemble of states, $\rho^{\otimes N}$, as the measurement of A is not only limited to the symmetric part of information about ρ, but it also affects the state by projecting it onto the symmetric subspace, thus changing to $\mathcal{G}_A(\rho)$. Therefore having a single copy of ρ and taking into account the no-cloning theorem [14], one cannot in general obtain all the information about ρ. One can formulate this problem in terms of the so-called error-disturbance relations, which quantitaively describe how the acquisition of information about one classical context (e.g. by measuring observable A) disturbs the information about the different classical context (e.g. connected with measurement outcomes of observable B). One of the ways to find such relations is to note that a given classical context of ρ connected with definite values of A is the symmetric component of ρ, $\mathcal{G}_A(\rho)$, and then to study the relations between symmetric subspaces with respect to different observables. In other words: if one wants to find how the measurement of A affects further measurements of B, on has to find out how asymmetric properties with respect to U(1) symmetry generated by B are transformed by projection onto the symmetric subspace of A. For example, if ρ is diagonal in the basis of eigenstates of A, so that it is symmetric with respect to A, then $\mathcal{G}_A(\rho) = \rho$ and the state, together with all its classical contexts, is not affected (disturbed). On the other hand if ρ is asymmetric with respect to A, then the measurement will change it, thus affecting some of the classical contexts of the state. This shows that the error-disturbance relations should be state-dependent, as they do not depend only on the symmetric subspaces of two observables, but on the projection of ρ onto them.

As mentioned above, the more asymmetric the state is with respect to A, the more information about it is missing if one is restricted only to the measurement outcomes (classical context) of A. To see this more clearly let us consider the following example. Let us imagine that Alice is given two copies of a pure state $|\psi_1\rangle$ and one of them undergoes an unknown symmetry operation, i.e.

$$|\psi_1\rangle \rightarrow |\psi_2\rangle = e^{i\phi A} |\psi_1\rangle , \tag{1.20}$$

for some unkown ϕ. Both states, $|\psi_1\rangle$ and $|\psi_2\rangle$, have the same classical contexts with respect to observable $A = \sum_n n |n\rangle \langle n|$,

$$\forall_n | \langle n|\psi_1\rangle |^2 = | \langle n|\psi_2\rangle |^2, \tag{1.21}$$

so if the only equipment in the Alice's lab is the device that measures A, then Alice cannot distinguish between $|\psi_1\rangle$ and $|\psi_2\rangle$. The first question is how much can the two states differ and still be indistinguishable by measurements of A, so how much information about a state is not encoded in the outcome probability distribution of A? As a measure of difference between two states $|\psi_1\rangle$ and $|\psi_2\rangle$ let us use the squared fidelity,

$$F^2(\psi_1, \psi_2) = | \langle \psi_1|\psi_2\rangle |^2. \tag{1.22}$$

Now it is easy to see that there exist states $|\psi_1\rangle$ such that the fidelity with the transformed state is minimal and equals 0, e.g. choose $|\psi_1\rangle = |0\rangle + |1\rangle$ and $\phi = \pi$, so that $|\psi_2\rangle = |0\rangle - |1\rangle$. Therefore it is more interesting to ask a second question: which states on average minimize the fidelity (differ most), where the average is taken over ϕ. This corresponds to a situation when Alice has two copies of a known state $|\psi_1\rangle$ and puts one into a black box that performs a symmetry operation with some random ϕ. The modified question is how much on average will the state from the black box, $|\psi_2\rangle$, differ from the original one (using the squared fidelity measure) and which state will differ most (so for which state the part of information encoded in the outcome probability distribution of A is the least). The average squared fidelity is given by

$$\begin{aligned} F^2_{avg}(\psi_1, \psi_2) &= \frac{1}{2\pi}\int_0^{2\pi} d\phi | \langle \psi_1|\psi_2\rangle |^2 = \frac{1}{2\pi}\int_0^{2\pi} d\phi \left| \sum_{m,n} \langle m|n\rangle \, a_m^* a_n e^{i\phi n} \right|^2 \\ &= \frac{1}{2\pi}\int_0^{2\pi} d\phi \left| \sum_n |a_n|^2 e^{i\phi n} \right|^2 = \frac{1}{2\pi}\int_0^{2\pi} d\phi \sum_{n,m} |a_n|^2 |a_m|^2 e^{i\phi(n-m)} \\ &= \sum_n |a_n|^4 \end{aligned}$$

and in order to find the state that minimizes this expression one can use the Lagrange multpliers method. Let us define the Lagrangian,

$$\mathcal{L} = \sum_n |a_n|^4 - \lambda \left(\sum_n |a_n|^2 - 1 \right) = \sum_n p_n^2 - \lambda \left(\sum_n p_n - 1 \right), \tag{1.23}$$

and minimize it,

$$\frac{\partial \mathcal{L}}{\partial p_m} = 2p_m - \lambda = 0, \tag{1.24}$$

thus obtaining $\forall_n p_n = -\lambda/2$, so together with the constraint: $\forall_n p_n = 1/N$, where N is the dimension of the Hilbert space $|\psi_1\rangle$ and $|\psi_2\rangle$ live in. Therefore a maximally asymmetric state (with respect to the previously introduced measures, the relative entropy of frameness and the variance) is also a state with the least information encoded in the symmetric subspace, i.e. a state for which the least information is available from the measurements of A.

Chapter 2

Sequential measurements

This chapter addresses the problem of sequential measurements of two noncommuting observables, in particular on the relation between the error (noise) ϵ of measuring the first observable A and the disturbance η (induced by the first measurement) of the measurement of the second observable B. In order to obtain a relation connecting these two quantities, one has first to find the physically meaningful definitions of both ϵ and η for quantum measurements of A and B. One of the possible requirements for such definitions is to make them dependent only on the difference between the probability distributions of a perfect measurement and the noisy or disturbed one. The physical justification and operative meaning of such choice is given and one of the possible definitions, based on the Kolmogorov variational distance, is proposed. Then the properties of noise and disturbance fulfilling the presented requirements are found and discussed for the simplest sequential measurement of two qubit observables (Pauli matrices along arbitrary directions). Next, the aproaches and definitions of noise and disturbance used by two groups (Ozawa's [2] and Busch's [9]) are presented and applied to the same problem of sequential measurement of qubit observables. It is then shown that in this case both approaches give the same result, which is inconsistent with the presented requirements for the noise and disturbance. It is pointed out that this inconsistency comes from the global, instead of state-dependent, definition of a perfect measurement. Other flaws of both approaches, independent of the proposed definitions of noise and disturbance, are also discussed showing that both Ozawa's and Busch's formulations fail to capture the error-disturbance relation for a quantum measurement of a system in a given initial state.

2.1 Proposal for the requirements for measurement error and disturbance

First of all let us point out that the noise and disturbance are not only functions of the measurements being performed (observables A and B), but also of the initial state ρ of the measured system. For given observables A and B there exist states that are more and less prone to error and disturbance, e.g. the measurement of σ_z will, in general, disturb the subsequent measurement of σ_x, however if the system is initially perpared in the eigenstate of σ_z no disturbance will be induced, as the state will not change after the measurement. Therefore we want to find the definitions of noise and disturbance dependent on the initial state of the measured system, and, as mentioned in Chapter 1, link them with the asymmetric properties with respect to $U(1)$ symmetries generated by measured observables A and B. One could argue that prior to the measurement one has no information on the state of the measured system, so that the proposed approach to find state-dependent definitions of noise and disturbance is unjustified. However having the state-dependent definitions allows to

- Find the worst and best case scenario, i.e. the state that maximizes or minimizes the noise, the disturbance or the noise-disturbance product.
- Average the noise-disturbance product over all possible initial states of the system.
- Although one cannot control the initial state of the measured system, one can control the initial state of the measuring probe, so that finding the state-dependent noise-disturbance relation could help in finding the best initial state of the probe that will minimize the noise-disturbance product.

Now let us focus on the meaning of the noisy and disturbed measurements. To be physically meaningful, the noisy (disturbed) measurement should be, in principle, distinguishable from the perfect one. It seems that the commonly accepted [2, 9] and natural concept of a perfect measurement of A is that the probability distribution of the measurement outcomes coincides with the theoretical probability distribution of A (sometimes it is called that such measurement satisfies the Born statistical formula). It is important to note that it is the theoretical probability distribution and not the probability amplitude that matters. Another condition that is often required from the perfect measurement is the so-called repeatability condition. It states that if the second measurement is performed immediately after the first one, then the two outcomes should be the same. Let us now discuss how these requirements for perfect measurement affect the definitions of the noise and disturbance.

In general, performing a projective measurement of some observable A on a state ρ will affect (disturb) it and change into $\rho' \neq \rho$ (apart from the special case, when ρ is diagonal in the basis of eigenstates of A). The same holds true for arbitrary POVM measurement $\mathcal{E}$. However, let us emphasize that we are interested in the disturbance of a subsequent measurement of B and not of a state. As was pointed out in Chapter 1, the projective measurement of observable B gives us only insight into the symmetric component of the measured state, $\mathcal{G}_B(\rho)$. Therefore the measurement will be affected only if the symmetric component of ρ is changed by the previous measurement of A (or POVM $\mathcal{E}$). Looking for a disturbance of the measurement of B we are therefore looking for the disturbance of one of the classical contexts of the quantum state ρ. As the symmetric components of ρ and ρ' (or equivalently classical contexts connected with observable B) are fully specified by the probability distributions, $p_n = \mathrm{Tr}\left(\rho \left|b_n\right\rangle \left\langle b_n\right|\right)$ and $p'_n = \mathrm{Tr}\left(\rho' \left|b_n\right\rangle \left\langle b_n\right|\right)$, the disturbance should only depend on them. The Born statistical formula for perfect (not disturbed) measurement is fulfilled when $\forall_n p_n = p'_n$, so the disturbance should be defined as a difference (defined in a suitable way) between these two probability distributions. The repeatability condition is fulfilled automatically, as the considered measurement of B is projective and therefore it collapses the measured state onto the eigenstate corresponding to the obtained outcome. We therefore conclude that, under the assumption that disturbed measurement should be distinguishable from the perfect one (either by failing to fulfill the Born statistical formula requirement or the repeatability condition), the disturbance should only depend on the difference between p_n and p'_n and vanish if $\forall_n p_n = p'_n$. The operative meaning of such definition is that the experimenter that is measuring observable B (whose knowledge is limited only to the measurement outcomes of B) should be able to distinguish between two ensembles of states, $\rho^{\otimes N}$ and $\rho'^{\otimes N}$, if the measurement of A (or POVM $\mathcal{E}$) disturbs the measurement of B.

Let us now consider the noise of the quantum measurement induced by performing the POVM $\mathcal{E}$ on a state ρ instead of the projective measurement of A. One can use similar arguments as in the disturbance case: the measurement of A gives insight only into symmetric part of ρ, $\mathcal{G}_A(\rho)$, which is fully specified by the probability distribution $q_n = \mathrm{Tr}\left(\rho \left|a_n\right\rangle \left\langle a_n\right|\right)$. Denoting the probability distribution of the POVM $\mathcal{E}$ by $q'_n = \mathrm{Tr}\left(M_n^\dagger \rho M_n\right)$, where M_n are detection operators of $\mathcal{E}$, the Born statistical formula for perfect (not noisy) measurement is fulfilled when $\forall_n q_n = q'_n$. The repeatability condition is in general not fulfilled, it is only if the support of ρ lies in the subspace spanned by orthogonal projectors of $\mathcal{E}$ (if $\mathcal{E}$ is a projective measurement then the repeatability condition is satisfied, since for projective measurements all projectors are orthogonal). This leaves us with two different notions of a noisy measurement. Firstly a measurement may

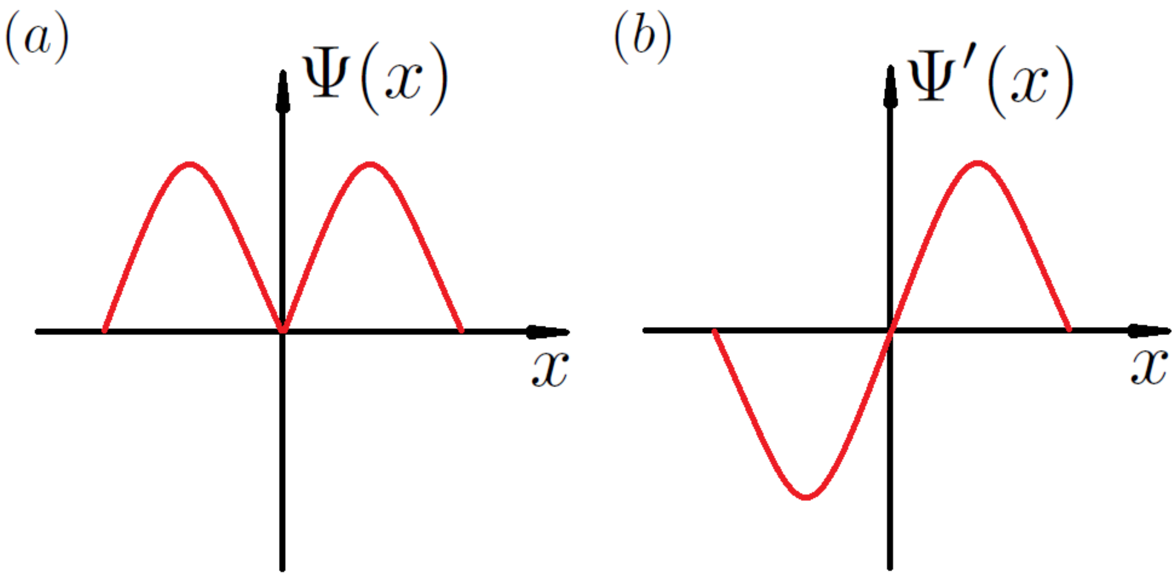

Figure 2.1: (a) Wavefunction before the POVM measurement $\mathcal{E}$; (b) Wavefunction after the POVM measurement $\mathcal{E}$.

be noisy, when it does not reproduce the statistics of the perfect measurement. Secondly it may be noisy if it fails to be repeatable. One can however argue, that the repeatability condition is not a neccessary requirement for the measurement to be perfect (not noisy). If one considers the measurement of A only as a way to get insight into the symmetric part of ρ, $\mathcal{G}_A(\rho)$, so as a process of gaining information about ρ and not also modifying it (by collapsing on the outcome eigenstate), then the only neccessary condition for a measurement to be perfect (not noisy) is to fulfill the Born statistical formula. We propose to use such approach, as it allows to define the noise only as a function of the difference between q_n and q'_n, similarly to the definition of disturbance, thus allowing for similar treatment of these two quantities and giving a well defined meaning of the noise-disturbance trade-off relation.

Let us now take a closer look to better understand the consequences of the fact that noise and disturbance depend only on the difference between actual and perfect measurement outcome statistics. For disturbance this means that if POVM measurement $\mathcal{E}$ disturbs the initial state of the system in such a way that it only changes the relative phases between eigenstates of B (which is equivalent to the action of U(1)-invariant unitary, see Eq. 1.12),

$$\rho = \sum_{m,n} \rho_{mn} \left| b_m \right\rangle \left\langle b_n \right| \xrightarrow{\text{POVM } \mathcal{E}} \rho' = \sum_{m,n} \rho'_{mn} \left| b_m \right\rangle \left\langle b_n \right|, \ \rho_{nn} = \rho'_{nn} \tag{2.1}$$

then the measurement of B is not disturbed. An example of such situation is presented in Fig. 2.1: before the POVM measurement $\mathcal{E}$ the wavefunction in position representation

$\Psi(x)$ is shown in panel (a), and the wavefunction $\Psi'(x)$ modified by this measurement is shown in panel (b). Although the wavefunction gets modified, the probability distribution of the position measurement is the same in both cases. Since it is impossible to detect this modification by measurement of x, the position measurements before and after the action of $\mathcal{E}$ are indistinguishable and therefore the disturbance should vanish. As another example consider a qubit prepared initially in the eigenstate of σ_z, $|0\rangle$. Now let σ_y be the first (disturbing) measurement and σ_x the second (disturbed) measurement. Even though the initial state $|0\rangle$ is highly disturbed by the measurement of σ_y (as it brings pure state to a maximally mixed one), the measurement of σ_x is not disturbed, as the two outcome statistics (with and without preceeding measurement of σ_y) are exactly the same, $p_+ = p_- = 1/2$. Similar example can be given for the case of the noisy measurement. Let the initial state of a qubit be again $|0\rangle$ and the perfect measurement one wants to perform be σ_x. Now any measurement along the direction lying in the XY plane of the Bloch sphere, e.g. σ_y, has vainishing noise, as it yields the same probability outcome distribution as the desired measurement of σ_x. This is a direct consequence of the fact that the symmetric components of $|0\rangle$ with respect to symmetries generated by both σ_x and σ_y are the same.

2.2 Error and disturbance based on the Kolmogorov variational distance

According to the reasoning presented in the previous section, the definitions of noise and disturbance should only be dependent on the suitably defined difference between the probability distributions of a perfect and actual measurement. Divergences, which are the functions that establish distance measures between probability distributions on a statistical manifold, are perfect candidates for such definitions. In this section we will propose the use of one of them, the Kolmogorov variational distance, justify this choice by giving its operational meaning and find the analytical expressions for noise and disturbance.

Let us start in a quantum information fashion by considering two games that will give us the operational intepretation of the proposed definitions. In the Game 1 (the quantum disturbance game) Alice is in possesion of an ensemble of identically prepared states, $\rho^{\otimes N}$, a device that performs a projective measurement of observable B and a black box. The action of the black box on an input state ρ is doing nothing (applying identity operation) with probability 1/2 and performing a POVM measurement $\mathcal{E}$ (but not giving the outcome to Alice), also with probability 1/2. In this game Alice has to put all her states into a black box and her goal is to guess what was the action of the black box on each state. If she did not have the device that measures B, her average success rate would be 1/2 (just blind guessing). Having this device however gives her access to the symmetric part of the measured state. Therefore if the symmetric part of ρ differs from the symmetric part of $\mathcal{E}(\rho)$, her average success rate will exceed 1/2. This increase in the success rate quantifies how sensitive the outcome probability distribution for the measurement of B is to the action of $\mathcal{E}$. Let us denote this probability distribution when black box does nothing by $p_n(\rho)$ and when it performs POVM $\mathcal{E}$ by $p'_n(\rho) = p_n(\mathcal{E}(\rho))$. Using the maximum likelihood estimation (MLE) method Alice's success rate for distinguishing between the two actions of a blackbox is given by

$$p_D = \frac{1}{2} + \frac{1}{4}\sum_n |p_n(\rho) - p'_n(\rho)| = \frac{1}{2} + K(p_n(\rho), p'_n(\rho)), \tag{2.2}$$

where $K(p_n(\rho), p'_n(\rho))$ is the Kolmogorov variational distance between probability distributions $p_n(\rho)$ and $p'_n(\rho)$. One can use p_D to quantify disturbance, as it measures the probability of distinguishing a perfect measurement from a disturbed one. In order to get vanishing disturbance for a perfect measurement and 1 for maximally disturbed one, the disturbance of a measurement of B in a state ρ caused by preceeding POVM measurement $\mathcal{E}$ can be defined as

$$\eta(B, \mathcal{E}, \rho) = 2K(p_n(\rho), p'_n(\rho)). \tag{2.3}$$

Game 2 (the quantum noise game) is similar to the first one. Alice is again in possesion of an ensemble $\rho^{\otimes N}$ and a different black box. This time the black box performs a projective measurement of observable A with probability 1/2 or a POVM measurement $\mathcal{E}$, also with probability 1/2. In both cases Alice is informed about the outcome. The goal of the game is again to guess which action the black box performed. Let $q_n(\rho)$ and $q'_n(\rho)$ be the probability distributions of the measurement outcomes of A and $\mathcal{E}$, respectively. Using MLE Alice's success rate is given by

$$p_N = \frac{1}{2} + K(q_n(\rho), q'_n(\rho)). \tag{2.4}$$

Similarly to the disturbance case the noise of the measurement of A in a state ρ when POVM $\mathcal{E}$ is performed instead, can be defined as

$$\epsilon(A, \mathcal{E}, \rho) = 2K(q_n(\rho), q'_n(\rho)). \tag{2.5}$$

One can express ϵ and η with the use of detection operators of $\mathcal{E}$, M_n, and the eigenstates of A and B, $\{|a_n\rangle\}$ and $\{|b_n\rangle\}$,

$$\epsilon(A, \mathcal{E}, \rho) = \frac{1}{2}\sum_n \left|\mathrm{Tr}\left(\rho(|a_n\rangle\langle a_n| - M_n^\dagger M_n)\right)\right|, \tag{2.6a}$$

$$\eta(B, \mathcal{E}, \rho) = \frac{1}{2}\sum_n \left|\mathrm{Tr}\left(\rho(|b_n\rangle\langle b_n| - \sum_m M_m |b_n\rangle\langle b_n| M_m^\dagger)\right)\right|. \tag{2.6b}$$

Unfortunately no error-disturbance trade-off relation has yet been found for the above definitions of ϵ and η.

It is also interesting to note that there exists no projective measurement that has disturbance equal to 1, i.e. that can always be detected in the scheme presented above. The optimal arrangement to detect a measurement of A by measuring B is to prepare the initial state in the eigenstate of B, $|b_0\rangle$ (in general the sharper the distribution, the easier it is to disturb it). Then the measurement of A transforms the state into

$$\rho' = \sum_n |\langle a_n|b_0\rangle|^2 |a_n\rangle\langle a_n|. \tag{2.7}$$

and the probability of getting b_0 outcome from measurement of B changes from 1 to

$$p = \sum_n |\langle a_n|b_0\rangle|^4 = \sum_n p_n^2. \tag{2.8}$$

The corresponding disturbance (scaled Kolmogorov distance) is then equal to

$$\eta(B, A, |b_0\rangle) = 2K = 1 - p. \tag{2.9}$$

In order to find the maximal possible disturbance over all measurements of A one has to find the extremum of p, which can be done using the Lagrange multipliers method with the constraint $\sum_n p_n = 1$. Defining Lagrangian

$$\mathcal{L} = \sum_n p_n^2 + \lambda\left(\sum_n p_n - 1\right), \tag{2.10}$$

one has

$$\frac{\partial\mathcal{L}}{\partial m} = \sum_n 2p_n\frac{\partial p_n}{\partial p_m} + \lambda\left(\sum_n \frac{\partial p_n}{\partial p_m}\right) = 2p_m + \lambda = 0, \tag{2.11}$$

so that $\forall_n p_n = -\lambda/2$, which, taking the constraint into account, gives $\forall_n p_n = 1/N$, where N is the dimension of the Hilbert space. Therefore the minimal p is

$$p_{min} = \sum_n p_n^2 = \sum_n \frac{1}{N^2} = \frac{1}{N}, \tag{2.12}$$

and the corresponding maximal possible disturbance

$$\eta_{max}(B, A, |b_0\rangle) = 1 - \frac{1}{N}. \tag{2.13}$$

2.3 Sequential measurements of a qubit

Let us now discuss the properties of noise and disturbance fulfilling the presented requirements for the simplest case of two qubit observables. The sequential measurement scenario is as follows. The system is initially prepared in a pure state

$$|\psi\rangle\langle\psi| = \frac{1 + \boldsymbol{c}\cdot\boldsymbol{\sigma}}{2}. \tag{2.14}$$

The perfect measurement that one wants to perform is given by

$$A = \underline{\boldsymbol{a}} \cdot \underline{\boldsymbol{\sigma}} = |a_0\rangle \langle a_0| - |a_1\rangle \langle a_1| \,, \tag{2.15}$$

the measurement that is actually performed (noisy one) is given by

$$M = \underline{\boldsymbol{m}} \cdot \underline{\boldsymbol{\sigma}} = |m_0\rangle \langle m_0| - |m_1\rangle \langle m_1| \,, \tag{2.16}$$

and the measurement that is disturbed by measuring M is given by

$$B = \underline{\boldsymbol{b}} \cdot \underline{\boldsymbol{\sigma}} = |b_0\rangle \langle b_0| - |b_1\rangle \langle b_1| \,. \tag{2.17}$$

As a result of a simple calculation one obtains the probability distributions of a perfect and noisy measurements of A, i.e. of A and M:

Perfect measurement	Noisy measurement
$p(A = a_0) = \frac{1+\underline{\boldsymbol{c}}\cdot\underline{\boldsymbol{a}}}{2}$	$p(M = m_0) = \frac{1+\underline{\boldsymbol{c}}\cdot\underline{\boldsymbol{m}}}{2}$
$p(A = a_1) = \frac{1-\underline{\boldsymbol{c}}\cdot\underline{\boldsymbol{a}}}{2}$	$p(M = m_1) = \frac{1-\underline{\boldsymbol{c}}\cdot\underline{\boldsymbol{m}}}{2}$

Similarly one can obtain the probability distributions of a perfect and disturbed (by measuring M) measurements of B:

Perfect measurement	Disturbed measurement
$p(B = b_0) = \frac{1+\underline{\boldsymbol{c}}\cdot\underline{\boldsymbol{b}}}{2}$	$p'(B = b_0) = \frac{1+\underline{\boldsymbol{c}}\cdot\underline{\boldsymbol{m}}\underline{\boldsymbol{m}}\cdot\underline{\boldsymbol{b}}}{2}$
$p(B = b_1) = \frac{1-\underline{\boldsymbol{c}}\cdot\underline{\boldsymbol{b}}}{2}$	$p'(B = b_1) = \frac{1-\underline{\boldsymbol{c}}\cdot\underline{\boldsymbol{m}}\underline{\boldsymbol{m}}\cdot\underline{\boldsymbol{b}}}{2}$

Let us first analyze the general properties of noise and disturbance that are independent of the precise definitions of ϵ and η (as long as they depend only on probability distributions of the outcomes):

1. If $\underline{\boldsymbol{c}} = \underline{\boldsymbol{m}}$, i.e. the initial state $|\psi\rangle$ is the eigenstate of the measured observable M, then there is no change in the statistics of the measurement outcomes of B and therefore no disturbance. This holds more generally for any system: if the initial state ρ is diagonal in the basis of eigenstates of the first measurement, then no measurement is disturbed. The reason for this is that in such case the initial state is not modified by the measurement, $\rho' = \rho$, so none of its classical contexts change.

2. If $\underline{\boldsymbol{c}} \perp \underline{\boldsymbol{m}}$, i.e. the initial state $|\psi\rangle$ is maximally asymmetric state of the measured observable M, then the statistics of the measurement outcomes of B is changed to $\left(\frac{1}{2}, \frac{1}{2}\right)$ (for all B). This holds more generally as maximally asymmetric states of some observable A are always mapped to maximally mixed states by the measurement of A.

3. If $\underline{\boldsymbol{m}} \perp \underline{\boldsymbol{b}}$, i.e. the eigenstates of the measured observable M are maximally asymmetric states of the observable B (the eigenstates of M and B form maximally unbiased bases, which are one of the subjects discussed in Chapter 4), then the statistics of the measurement outcomes of B is changed to $\left(\frac{1}{2}, \frac{1}{2}\right)$ (for all $|\psi\rangle$). This holds more generally: for any 2 observables A and B if their eigenstates form mutually unbiased bases, then, independently of the system state ρ, the measurement of A makes the symmetric part of ρ with respect to B a maximally mixed state.

4. If for eigenstates of all observables A, B, M there exists a common completely asymmetric state (which means that the Bloch vectors $\underline{\boldsymbol{a}}$, $\underline{\boldsymbol{b}}$, $\underline{\boldsymbol{m}}$ lie in one plane) and the initial state is that state, then perfect probability distributions are the same as noisy and disturbed ones, so there is no noise and no disturbance. There is also a generalization of this fact, which will be discussed in more detail in Sec. 4.4. In short: for every dimension of Hilbert space there exist non-commuting observables A and B and a pure state $|\psi\rangle$ such that perfect sequential measurements (with vanishing noise and disturbance) of A and B are possible.

Another, however less rigorously defined, property is that the sharper the distribution of the initial pure state over the eigenstates of the observable A one wants to measure (the more symmetric the initial pure state is), the more prone the measurement is to errors. As an example consider that one wants to perform a measurement of $A = \sigma_z$ and that initially the system is in the symmetric state $|0\rangle$. Any deviation from measuring along z direction (by measuring observable $M = \underline{\boldsymbol{m}} \cdot \underline{\boldsymbol{\sigma}}$, $\underline{\boldsymbol{m}} \neq \hat{z}$) will affect the measurement outcomes statistics [change it from $(1, 0)$ to $(\frac{1+m_z}{2}, \frac{1-m_z}{2})$]. On the other hand, if the system is initially in the completely asymmetric state of $A = \sigma_z$, let say $|+\rangle$, then instead of measuring exactly along z direction one can perform any measurement in the YZ plane of the Bloch sphere and get the same outcome statistics, so there is no measurement error.

Finally let us find the Kolmogorov distance based expressions for the noise and disturbance of sequential measurements of a qubit

$$\epsilon = \frac{|\underline{\boldsymbol{c}} \cdot (\underline{\boldsymbol{a}} - \underline{\boldsymbol{m}})|}{2}, \tag{2.18}$$

$$\eta = \frac{|\underline{\boldsymbol{c}} \cdot (\underline{\boldsymbol{b}} - \underline{\boldsymbol{m}} \cdot \underline{\boldsymbol{b}}\ \underline{\boldsymbol{m}})|}{2}. \tag{2.19}$$

2.4 Ozawa's approach

In recent years the approach to the problem of the measurement noise and disturbance that got the most attention and recognition is the one proposed by Ozawa [2]. In this section we will first give a brief overview of the formalism he uses, his defintions of noise and disturbance and the resulting universally valid reformulation of the Heisenberg uncertainty principle [2]. Next, using Ozawa's approach, we will find the noise and disturbance of the previously discussed sequential measurements of qubit observables. Basing on this example we will point out the general flaws of Ozawa's formulation and its failure to capture the state-dependent noise and disturbance of quantum measurements.

2.4.1 Overview of the formalism

The approach used by Ozawa bases on the indirect measuring scheme. Instead of measuring the system observable A directly the system undergoes joint evolution U with the probe and then the probe observable M is measured. The measuring interaction U and the probe observable M are chosen so that the outcome statistics of the probe measurement coincides with the outcome statistics of the original measurement of A. Formally the system is prepared in the initial state $|\psi\rangle$ and its observables are given by

$$A = A \otimes \mathbb{1}, \tag{2.20a}$$

$$B = B \otimes \mathbb{1}. \tag{2.20b}$$

The measuring probe (meter) is prepared in the state $|\zeta\rangle$ and its observable is given by

$$M = \mathbb{1} \otimes M. \tag{2.21}$$

Then the composed system undergoes joint unitary evolution and the description is given in the Heisenberg picture, so that states of the system and probe are constant and the observables evolve into

$$A' = U^{\dagger}(A \otimes \mathbb{1})U, \tag{2.22a}$$

$$B' = U^{\dagger}(B \otimes \mathbb{1})U, \tag{2.22b}$$

$$M' = U^{\dagger}(\mathbb{1} \otimes M)U. \tag{2.22c}$$

The advantage of the indirect measurement scheme comes from the fact that, because $[B, M] = 0$ also $[B', M'] = 0$ and the observables B' and M' can be simultaneously measured. Therefore the problem of not simultaneously measureable observables A and B is changed into the problem of finding optimal U, M and $|\zeta\rangle$ such that M' simulates the measurement of A and B' simulates the measurement of B as precisely as possible.

The noise $N(A)$ and disturbance $D(B)$ operators are defined as the difference between the original and modified observables,

$$N(A) = M' - A, \tag{2.23a}$$

$$D(B) = B' - B, \tag{2.23b}$$

and the noise and disturbance are quantified by the root-mean-square of the above operators it the state $|\psi,\zeta\rangle$,

$$\begin{aligned}\epsilon(A) &= \langle N^2(A)\rangle^{\frac{1}{2}} = ||(U^\dagger(1\otimes M)U - A\otimes 1)\,|\psi,\zeta\rangle\,||, &(2.24a)\\ \eta(B) &= \langle D^2(B)\rangle^{\frac{1}{2}} = ||(U^\dagger(B\otimes 1)U - B\otimes 1)\,|\psi,\zeta\rangle\,||. &(2.24b)\end{aligned}$$

Ozawa also presented the reformulation of his approach in terms of of POVM measurements [15], which was later used in the paper on experimental verification of the theory [16]. The detection operators M_m for the POVM measurement are defined by

$$M_m\,|\psi\rangle = \langle m|\,U\,|\psi,\zeta\rangle\,,$$

or equivalently by

$$U\,|\psi,\zeta\rangle = \sum_m M_m\,|\psi,m\rangle\,,$$

where $|m\rangle$ are the eigenstates of the meter observable M with eigenvalues m. The definitions of noise and disturbance expressed with the use of these detection operators are given by

$$\begin{aligned}\epsilon^2(A) &= \sum_m ||M_m(m-A)\,|\psi\rangle\,||^2, &(2.25a)\\ \eta^2(B) &= \sum_m ||[M_m,B]\,|\psi\rangle\,||^2. &(2.25b)\end{aligned}$$

In his seminal paper [2] Ozawa proves the following inequality

$$\epsilon(A)\eta(B) + \epsilon(A)\sigma(B) + \sigma(A)\eta(B) \geq \frac{1}{2}|\langle[A,B]\rangle|, \tag{2.26}$$

where $\sigma(A) = \sqrt{\langle A^2\rangle - \langle A\rangle^2}$ is the standard deviation of A in the state $|\psi\rangle$. The above inequality gives the trade-off relation for the noise and disturbance of sequential quantum measurements and shows that it is dependent on the uncertainty of the measured operators in the initial state of the system. The sketch of the derivation of this inequality and ways to improve the bound will be presented in Sec. 4.1.3.

2.4.2 Application to sequential measurements of a qubit

Let us now apply Ozawa's definitions to the measurements of two arbitrary observables of a qubit (this is what was actually done in the paper on experimental verification of Ozawa bound using neutron spins [16]). The system observables, similarly as before, are given by

$$\begin{aligned}A &= \underline{\boldsymbol{a}}\cdot\underline{\boldsymbol{\sigma}}, &(2.27)\\ B &= \underline{\boldsymbol{b}}\cdot\underline{\boldsymbol{\sigma}}, &(2.28)\end{aligned}$$

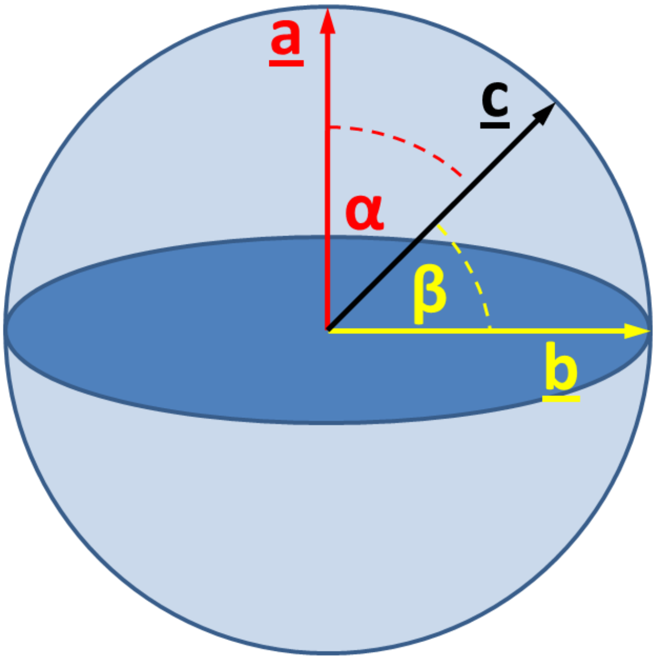

Figure 2.2: The definitions of angles appearing in Ozawa's expressions for error and disturbance of the qubit measurements visualised on the Bloch sphere.

and the indirect measurement (POVM which is actually a projective measurement) of observable C,

$$C = \underline{\boldsymbol{c}} \cdot \underline{\boldsymbol{\sigma}}, \tag{2.29}$$

is performed. There are infinitely many realizations of such measurement, i.e. there are infinitely many choices of $|\zeta\rangle$, M and U. Let us quickly describe one of such realizations, when the probe itself is also qubit: choose the meter observable to be $M = \sigma_z$, the initial state of the apparatus to be $|\zeta\rangle = |0\rangle$ and the joint unitary evolution to be $U = |c_0\rangle \langle c_0| \otimes \mathbb{1} + |c_1\rangle \langle c_1| \otimes \sigma_x$, where $|c_{0/1}\rangle$ denotes the positive/negative eigenstate of C. Then one gets the following measurement operators

$$\begin{aligned} M_0 |\psi\rangle &= \langle c_0|\psi\rangle |c_0\rangle, \\ M_1 |\psi\rangle &= \langle c_1|\psi\rangle |c_1\rangle, \end{aligned}$$

and the associated probabilities are $p_0 = |\langle c_0|\psi\rangle|^2$ and $p_1 = |\langle c_1|\psi\rangle|^2$, so that the above indirect measurement process perfectly simulates the measurement of C.

Simple calculations lead to the following result for the noise and disturbance for the measurements of the qubit observables

$$\epsilon(A) = 2\left|\sin\frac{\alpha}{2}\right|, \tag{2.30}$$

$$\eta(B) = \sqrt{2}\,|\sin\beta|, \tag{2.31}$$

where the angles α and β are defined as

$$\cos\alpha = \underline{\boldsymbol{a}} \cdot \underline{\boldsymbol{c}}, \tag{2.32}$$

$$\sin\beta = \underline{\boldsymbol{a}} \cdot \underline{\boldsymbol{b}}, \tag{2.33}$$

which can be visualised by looking at Fig. 2.2.

2.4.3 Discussion and criticism

Let us first focus on the properties of Ozawa's definitions of noise and disturbance for the case of sequential measurements of a qubit presented in the previous subsection. The most important feature is that the noise and disturbance are completely independent of the inital state of the measured system. Therefore they do not depend on the difference between probability distributions of the perfect and actual measurement, which was our requirement presented in this chapter. One could argue that Ozawa's treatment of the quantum measurement is different than the one proposed by us and that he may treat a measurement to be disturbed not only when the probability distribution of the outcomes changes, but more generally, when the state changes (something that we called the disturbance of a state, not of a measurement). However, even in such situation Ozawa's definition of disturbance fails to be valid (physically meaningful), and the easiest way to see this is to consider the following example. Let the system be initially in the eigenstate $|0\rangle$ of the measured observable $A = \sigma_z$ and let the second observable be $B = \sigma_x$. The measurement of a system that already is in an eigenstate of the measured observable does not change the system state at all, so it cannot disturb the subsequent measurement of B in any way. However the disturbance obtained from the Ozawa's expression in this case is maximum and equal to $\sqrt{2}$. The conclusion is that, in the case of qubit measurements, Ozawa's definitions of noise and disturbance not only fail to fulfill the very basic and physically motivated requirements presented in this chapter, but also give nonzero disturbance when there is no change in the system state.

Let us now proceed to the general discussion on the origin of the presented flaws of Ozawa's formulation. In his paper [2] Ozawa emphasizes the importance of a state-dependent definitions of the noise and disturbance and defines ϵ and η to be, seemingly, dependent on the state of the measured system. He also requires perfect measurement to satisfy the Born statistical formula. His justification of the definitions of ϵ and η is based on two theorems. The first one states: "The noise of the measurement of A, $\epsilon(A)$, is zero for all state $|\psi\rangle$ if and only if M' satisfies the Born statistical formula for all states $|\psi\rangle$ (i.e. the outcome statistics of M' is the same as of A for all states)". The second theorem states: "The disturbance of the measurement of B, $\eta(B)$, is zero for all state $|\psi\rangle$ if and only if B' satisfies the Born statistical formula for all states $|\psi\rangle$ (i.e. the outcome statistics of B' is the same as of B for all states)". As can be seen from the example of the qubit measurements this is not enough and, if one wants to have truly state-dependent definitions of noise and disturbance, one should actually demand stronger requirements. The requirements for the definition of noise and disturbance should be: for every state $|\psi\rangle$ for which the Born statistical formula is satisfied the noise ϵ (the disturbance η) should vanish. Interestingly the followers of Ozawa's approach in their paper criticizing alternative Busch's approach [17] write: "Busch *et al.*'s quantities describe the disturbing power of a measuring device, quantifying how much the measurement could disturb some hypothetical state, whereas Ozawa describes how much a given quantum state is disturbed". This proves that the general impression and understanding is that Ozawa's definition capture the state-dependent noise and disturbance, even though this is not true, as explained above.

2.5 Busch's approach

Let us now shortly describe the alternative approach proposed by Busch *et al.* [9]. For a given system state ρ and a sharp position value ζ Busch defines

$$D(\rho, Q'; \zeta) = \langle (q' - \zeta)^2 \rangle^{\frac{1}{2}}_{\rho, Q'}, \tag{2.34}$$

where q' are the results of approximate measurement Q' of position observable Q. Now the error of this approximate measurement is defined as

$$\Delta(Q, Q') = \lim_{\epsilon \to 0} \sup\{D(\rho, Q'; \zeta)|_{\rho,\zeta}; D(\rho, Q; \zeta) < \epsilon\}. \tag{2.35}$$

In other words: for every sharp value of position ζ find the state ρ, such that perfect measurement Q gives outcome ζ with error smaller than ϵ, i.e. find the best representation of a sharp value state. Then, for this state and sharp value, calculate D (root-mean-square difference) for the approximate measurement Q'. The error is maximized over all possible sharp value states (as, for example, an approximate measurement may be perfect in some region $x_1 < x < x_2$, but not perfect outside and in this case to find the measure of error one has to consider the worst case scenario). The disturbance of observable B is defined in the same way as $\Delta(B, B')$, where B' is the effective measurement that one performs when disturbing measurement of A was first done.

Let us now apply the above definitions for sequential measurements of a qubit. We will use the notation already used in Sec. 2.3, Eqs. 2.14-2.17. In the discrete case one does not have to take the limit $\epsilon \to 0$ as there exist correct density matrices representing sharply valued states: for the first observable one wants to measure, A, we have

$$\begin{aligned}\zeta_0 = 1 \quad &\rightarrow \quad \rho_0 = \frac{1 + \boldsymbol{a}\cdot\boldsymbol{\sigma}}{2} \\ \zeta_1 = -1 \quad &\rightarrow \quad \rho_1 = \frac{1 - \boldsymbol{a}\cdot\boldsymbol{\sigma}}{2},\end{aligned}$$

and for the second (disturbed) one, B,

$$\begin{aligned}\zeta_0' = 1 \quad &\rightarrow \quad \rho_0' = \frac{1 + \boldsymbol{b}\cdot\boldsymbol{\sigma}}{2} \\ \zeta_1' = -1 \quad &\rightarrow \quad \rho_1' = \frac{1 - \boldsymbol{b}\cdot\boldsymbol{\sigma}}{2}.\end{aligned}$$

Now, using the definition of Δ one finds that the error and noise are exactly the same as in the Ozawa's case,

$$\begin{aligned}\Delta(A, M) &= 2\left|\sin\frac{\alpha}{2}\right| = \epsilon(A), \\ \Delta(B, M) &= \sqrt{2}|\sin\beta| = \eta(B),\end{aligned}$$

with angles α and β already defined by Eqs. 2.32-2.33.

The definitions of error and disturbance defined in the way proposed by Busch (in general, not only in the case of qubits) are only good for initial state of the system being the eigenstate of the observable one wants to measure - they quantify how the sharply valued states are prone to error or disturbance. This is the upper bound of the error (disturbance) taken over all initial states of the measured system, as eigenstates are most prone to error (disturbance), as mentioned in Sec. 2.3.

Looking for the trade-off relation between the error of measuring A and disturbance of measuring B does not have much sense with the Busch's definitions of ϵ and η. This is because $\epsilon(A)$ is defined for initial state being an eigenstate of A and the disturbance of B is defined for initial state being an eigenestate of B and both this conditions cannot be fulfilled simultaneously if $[A, B] \neq 0$. The product of the error of A and disturbance of B according to Busch's definitions has the following sense: the maximum error of A (over all possible system states) times the maximum disturbance of B (over all possible system states), whereas it should be: [the maximum error times the maximum disturbance] (over all possible states). In other words, the maximization should be done simultaneously, not separately. This is the point that also the Ozawa's followers did in their paper criticizing Busch [17]. In conclusion, we find that the Busch's approach does not capture the state-dependent problem of the noise and disturbance of quantum measurements.

Chapter 3

Measurements in the presence of a conservation law

This chapter presents the problem of quantum measurements in the presence of a conservation law, i.e. the limitation that a conservation law imposes on the information available about the quantum state through measurements. The historical overview of the traditional approaches to the measurement of quantum mechanical operators will be given, starting from the Wigner's seminal work [3], through the formal definition of the Wigner-Araki-Yanase (WAY) theorem by Araki and Yanase [4, 18] to the most recent result for the upperbound of the measurement precision by Ozawa [6]. Then the formulation of the WAY theorem in terms of asymmetry resources will be given, and its proof will be presented with the use of the tools introduced in Chapter 1.

3.1 Historical overview

In 1952 Wigner considered a simple measurement scheme in the presence of a conservation law and showed that if the measured observable does not commute with the conserved quantity, the repeatable measurement is impossible [3]. Specifically, he focused on the measurement of the spin component along x direction, when the total angular momentum (of the system and apparatus) along z is conserved. He used the general definition of the measurement process of observable A in terms of a joint unitary evolution U_A of the measured system, initially in the state $|\phi\rangle$, coupled to measuring apparatus, prepared in the state $|\xi\rangle$,

$$U_A(|\phi\rangle \otimes |\xi\rangle) = \sum_n \langle a_n|\phi\rangle\, |a_n\rangle \otimes |\chi_n\rangle\,, \tag{3.1}$$

where $|a_n\rangle$ are the eigenstates of A and demanding that the appartus states $|\chi_n\rangle$ are macroscopically distinguishable, i.e. $\langle\chi_n|\chi_m\rangle = \delta_{mn}$. Therefore the pointer (apparatus) state $|\chi_n\rangle$ indicates the measurement outcome a_n, which is the eigenvalue corresponding to the eigenstate $|a_n\rangle$. Note that this measurement process satisfies both the Born statistical formula (as measurement outcome $|a_n\rangle$ is obtained with probability $|\langle a_n|\phi\rangle|^2$) and the repeatibility condition (as obtaining the measurement outcome a_n collapses the system state to $|a_n\rangle$).

Now, applying the introduced measuring scheme to a particular case of spin 1/2 measurement along the x direction, one gets

$$U_{S_x}(|+\rangle|\xi\rangle) = |+\rangle|\chi\rangle, \tag{3.2a}$$

$$U_{S_x}(|-\rangle|\xi\rangle) = |-\rangle|\chi'\rangle, \tag{3.2b}$$

where $|\pm\rangle = (|0\rangle \pm |1\rangle)/\sqrt{2}$ are the eigenstates of the measured S_x spin operator. To obtain the effect of measurement when the system is initially in the eigenstate of S_z operator, one simply adds and subtracts the above equations,

$$U_{S_x}(|0\rangle|\xi\rangle) = |0\rangle|\sigma\rangle + |1\rangle|\tau\rangle, \tag{3.3a}$$

$$U_{S_x}(|1\rangle|\xi\rangle) = |0\rangle|\tau\rangle + |1\rangle|\sigma\rangle, \tag{3.3b}$$

where $|\sigma\rangle = (|\chi\rangle + |\chi'\rangle)/\sqrt{2}$ and $|\tau\rangle = (|\chi\rangle - |\chi'\rangle)/\sqrt{2}$. Now it is easily seen that if the total angular momentum along z, L_z, has to be conserved, such an evolution is impossible. This is because the expectation value of L_z on the right hand sides of Eqs. 3.3a-3.3b is the same, whereas on the left hand sides it differs by 1/2.

As the repeatable measuring scheme introduced in Eq. 3.1 fails to be universal and, as Wigner argued, "a measurement of spin components is practically possible", an approximate measuring scheme must exist. Therefore first he described a measuring process as a general unitary evolution of the joint system,

$$U_{S_x}(|+\rangle|\xi\rangle) = |+\rangle|\chi\rangle + |-\rangle|\eta\rangle, \tag{3.4a}$$

$$U_{S_x}(|-\rangle|\xi\rangle) = |-\rangle|\chi'\rangle + |+\rangle|\eta'\rangle, \tag{3.4b}$$

and then argued that if the norms of the unnormalized pointer states $|\eta\rangle$ and $|\eta'\rangle$ can be made arbitrarily small and $\langle\chi|\chi'\rangle = 0$, then the repeatable measurement can be recovered with arbitrary precision. By a suitable choice of the pointer states he then showed that this precision is limited by

$$\langle\eta|\eta\rangle = \frac{1}{2n-1}, \tag{3.5}$$

where n is the upper bound for the angular momentum along z of the measuring apparatus. Hence he proved that the precision of measurement is limited by the appropriately defined size of the measuring apparatus.

A few years after the Wigner's paper was published, Araki and Yanase generalized and formalized his work by stating and proving what is now known as the WAY theorem

[4]. They considered a measuring process described by Eq. 3.1 in a slightly more general way by including possible degeneracies of the measured observable A. Here, however, the degenerate spectrum will not be considered, as it does not change the reasoning significantly, but makes the notation much more unclear (for details see [4]). The measurement is thus described by

$$U_A(|a_n\rangle \otimes |\xi\rangle) = |a_n\rangle \otimes |\chi_n\rangle \,. \tag{3.6}$$

The conservation law is formulated in terms of additive, self-adjoint operator $L = L_1 \otimes \mathbb{1} + \mathbb{1} \otimes L_2$ commuting with any unitary time-evolution operator U,

$$[U, L] = 0, \tag{3.7}$$

which in particular includes the measurement process $U = U_A$. Now, from the additivity of L one has

$$\begin{aligned} \langle a_n \otimes \xi | L(a_m \otimes \xi)\rangle &= \langle a_n|a_m\rangle \langle \xi|L_2\xi\rangle + \langle a_n|L_1 a_m\rangle \langle \xi|\xi\rangle \\ &= \delta_{mn} \langle \xi|L_2\xi\rangle + \langle a_n|L_1 a_m\rangle \,. \end{aligned} \tag{3.8}$$

On the other hand, form unitarity of U_A and commutation relation given by Eq. 3.7 one also has

$$\begin{aligned} \langle a_n \otimes \xi | L(a_m \otimes \xi)\rangle &= \langle U_A(a_n \otimes \xi) | U_A L(a_m \otimes \xi)\rangle \\ &= \langle U_A(a_n \otimes \xi) | L U_A(a_m \otimes \xi)\rangle \\ &= \langle a_n \otimes \chi_n | L(a_m \otimes \chi_m)\rangle \\ &= \langle a_n|L_1 a_m\rangle \langle \chi_n|\chi_m\rangle + \langle a_n|a_m\rangle \langle \chi_n|L_2\chi_m\rangle \\ &= \delta_{mn}(\langle a_n|L_1 a_m\rangle + \langle \chi_n|L_2\chi_m\rangle). \end{aligned} \tag{3.9}$$

Thus, by comparing Eqs. 3.8 and 3.9, the following condition is obtained,

$$\langle a_n|L_1 a_m\rangle = \delta_{mn} \langle a_n|L_1 a_m\rangle \,. \tag{3.10}$$

This condition however implies that $[L_1, A] = 0$, because

$$\begin{aligned} \langle a_n|A L_1 a_m\rangle &= \lambda_n \langle a_n|L_1 a_m\rangle = \lambda_n \delta_{mn} \langle a_n|L_1 a_m\rangle \,, \\ \langle a_n|L_1 A a_m\rangle &= \lambda_m \langle a_n|L_1 a_m\rangle = \lambda_n \delta_{mn} \langle a_n|L_1 a_m\rangle \,. \end{aligned}$$

Therefore Araki and Yanase concluded that for the measurement process described by Eq. 3.6 to be possible in the presence of the additive conservation law, the measured observable must commute with the conserved quantity.

In his following work Yanase focused on the possibilities and limits of the approximate measurement schemes. He pointed out an important requirement, which is now known as the Yanase condition, that the pointer observable should commute with the conserved quantity. Otherwise, as he argued, the problem of precise measurement would just be shifted from the system to the measuring apparatus, as the WAY theorem also applies to the apparatus observable. He limited his considerations to the spin 1/2 measurement along the x direction and used the general measurement scheme proposed by Wigner and given by Eqs. 3.4a-3.4b. He also defined the lower bound for the probability of the unsuccesful measurement by the minimum of

$$\epsilon = \langle\eta|\eta\rangle + \langle\eta'|\eta'\rangle\,, \tag{3.11}$$

and introduced the measure of the apparatus size

$$M^2 = \langle\xi|L\xi\rangle\,, \tag{3.12}$$

i.e. the mean square value of the conserved quantity in the initial state of the apparatus. With the use of a few approximations Yanase obtained the following bound for ϵ,

$$\epsilon \geq \frac{1}{M^2}, \tag{3.13}$$

which is again connected with the appropriately defined size of the measuring apparatus.

The recent works of Loveridge and Busch extend the results presented above. Firstly, they showed that the original Araki and Yanase statement about the impossibility of a repeatable measurement of A when $[A, L_1] \neq 0$ can be reformulated in the following way: if $[A, L_1] \neq 0$ then the measurement of A must violate both the repeatability and the Yanase condition [19]. This means that not only the repeatable measurements are impossible but also the precise ones. Secondly, note that all results presented in this review are restricted to observables with discrete spectrum. Loveridge and Busch extended these considerations to continous unbounded quantities and provided the evidence for a WAY-type theorem [20].

3.2 Ozawa's approach

Quite recently a new approach to find the precision limits of approximate measurements was proposed by Ozawa [6]. Instead of devising the unitary process responsible for the measurement, he used the generalized Heisenberg uncertainty relation and obtained tight limit independent of the chosen measuring scheme. His description of the measuring process and the formalism he uses is very similar to the one presented in Sec. 2.4. However, as his work on the WAY theorem preceeded his results on the error-disturbance relation, we will present his original approach to provide the historical context and, at the end of this section, we will show how the obtained bound for the precision of the WAY scenario measurement follows directly from the application of the error-disturbance relation to a special case.

Initially the measured system **S** is prepared in an unknown state $|\psi\rangle$ and the probe (apparatus) **P** is in the known state $|\xi\rangle$, so that the total system is described by $|\psi\rangle \otimes |\xi\rangle$. In order to measure the observable A of the system **S**, the system is coupled to the apparatus and the composite system $\mathbf{S} + \mathbf{P}$ undergoes unitary evolution $U(t)$ during the measurement time t. After the measuring interaction the probe observable M is measured and the outcome is recorded as the value of macroscopic outcome variable $\mathbf{x}$. Using the Heisenberg picture, i.e. $A(t) = U^\dagger(t)(A \otimes \mathbb{1})U(t)$ and $M(t) = U^\dagger(t)(\mathbb{1} \otimes M)U(t)$, the probability distribution for $\mathbf{x}$ is given by

$$\Pr(\mathbf{x} \in \Delta) = ||E^{M(t)}(\Delta)(|\psi\rangle \otimes |\xi\rangle)||^2, \tag{3.14}$$

where $E^{M(t)}(\Delta)$ is the spectral projection of the operator $M(t)$ corresponding to the interval Δ.

Now, according to Ozawa's definition, the measurement is said to be precise if for all states the probability distribution of the macroscopic outcome variable $\mathbf{x}$ is the same as the probability distribution given by the Born statistical formula, i.e.

$$\forall\, |\psi\rangle : \Pr(\mathbf{x} \in \Delta) = ||E^{A(0)}(\Delta)\, |\psi\rangle\, ||^2. \tag{3.15}$$

To measure the departure from the precise measurement, the noise operator N is introduced,

$$N = M(t) - A(0), \tag{3.16}$$

which measures the difference between the obtained outcome and the actual state of the system just before the measurement. The noise $\epsilon(|\psi\rangle)$ of the measurement for a certain initial state $|\psi\rangle$ of the measured system is defined by the root-mean-square error in the outcome of the measurement,

$$\epsilon^2(|\psi\rangle) = \langle N^2\rangle \geq (\Delta N)^2, \tag{3.17}$$

where $(\Delta N)^2 = \langle N^2\rangle - \langle N\rangle^2$ denotes the variance of the nosie observable. The precision limit ϵ for arbitrary initial state is the maximum of $\epsilon(|\psi\rangle)$ taken over all possible states of the measured system **S**. The above definition of the noise is justified by the following theorem: measurement is precise if and only if $\epsilon = 0$.

Now one can find the tight bound for the measurement precision by using the general uncertainty relation for the noise operator N and the conserved quantity $L(0) = L_1(0) \otimes \mathbb{1} + \mathbb{1} \otimes L_2(0)$,

$$(\Delta N)^2(\Delta L(0))^2 \geq \frac{1}{4}|\langle[N, L(0)]\rangle|^2. \tag{3.18}$$

Employing the fact that $L_1(0)$ and $L_2(0)$ are statistically independent one gets $[\Delta L(0)]^2 = [\Delta L_1(0)]^2 + [\Delta L_2(0)]^2$. Also, since the measured system and probe operators commute, one has $[M(t), L_1(t)] = [A(0), L_2(0)] = 0$. Taking all these into account one obtains the following bound for $\epsilon(|\psi\rangle)$,

$$\epsilon^2(|\psi\rangle) \geq \frac{|\langle [M(t), L_2(t)] - [A(0), L_1(0)]\rangle|^2}{4[\Delta L_1(0)]^2 + 4[\Delta L_2(0)]^2}. \tag{3.19}$$

One could think, looking at the above expression, that by choosing $|\xi\rangle$, $U(t)$ and M so that

$$\langle\xi| U^\dagger(t)(\mathbb{1} \otimes [M, L_2])U(t) |\xi\rangle = [A, L_1], \tag{3.20}$$

the measurement noise ϵ can be put to zero. Note however that without the Yanase condition fulfilled, one cannot measure M precisely and another measuring apparatus and noise operator should be introduced to measure the probe system **P**. Such a sequence of measuring apparatuses and probe observables is necessary until one of them will commute with the conserved quantity. Therefore, effectively, one can consider only the case for which the Yanase condition is fulfilled, i.e. $[M, L_2] = 0$. Hence the fundamental noise bound is given by

$$\epsilon^2(|\psi\rangle) \geq \frac{|[A(0), L_1(0)]\rangle|^2}{4[\Delta L_1(0)]^2 + 4[\Delta L_2(0)]^2}. \tag{3.21}$$

Let us now show, how the above bound can be obtained with the use of Ozawa's universally valid error-disturbance uncertainty relation, given by Eq. 2.26. It is enough to choose the second observable B to be the additive conserved quantity,

$$B = L = L_1 \otimes \mathbb{1} + \mathbb{1} \otimes L_2, \tag{3.22}$$

and note the the conservation law enforces the disturbance $\eta(L)$ to vanish. The error-disturbance relation then changes to

$$\epsilon(A)\sigma(L) \geq \frac{1}{2}|\langle [A, L_1]\rangle|, \tag{3.23}$$

which can be easily transformed to Eq. 3.21.

3.3 Asymmetry resource approach

Now, basing on the Ref. [10], the proof of the WAY theorem, with the use of the resource theory of asymmetry formalism introduced in Chapter 1, will be presented.

The joint time evolution of the measured system and measuring apparatus is described by the unitary operator $U(t)$. The presence of the additive conservation law for operator $N = N_1 \otimes \mathbb{1} + \mathbb{1} \otimes N_2$ restricts this time evolution to unitaries $U(t)$ invariant under transformations generated by N, so that

$$[U(t), N] = 0. \tag{3.24}$$

One can show that the restriction induced by this conservation law for the quantum operation $\mathcal{E}$, describing the time evolution of the system alone, is that $\mathcal{E}$ must be U(1)-covariant [10]. Therefore to describe the measurement process one can focus only on the transformations of the measured system (omitting the apparatus) that are U(1)-covariant.

In order to precisely measure the observable A with non-degenerate spectrum, $A = \sum_i a_i |a_i\rangle\langle a_i| = \sum_i a_i\rho_i$, and in the presence of a conservation law for N one requires POVM quantum operation $\mathcal{M} = \sum_i \mathcal{M}_i$ such that

$$\begin{aligned} \forall i \neq k : \mathcal{M}_i(\rho_k) &= 0, & (3.25a) \\ \forall \rho : \sum_i \mathrm{Tr}[\mathcal{M}_i(\rho)] &= 1, & (3.25b) \\ \forall g \in G : [\mathcal{M}, \mathcal{U}(g)] &= 0. & (3.25c) \end{aligned}$$

In other words one needs to find U(1)-covariant quantum operation $\mathcal{M}$ that allows for perfect unambiguous discrimination between the eigenstates of A (perfect meaning that the probability of measurement failure is zero, i.e. every measurement is succesful). This can be achieved by considering all possible quantum operations without the restriction to U(1)-covariant transformations, which are then projected onto the set of the U(1)-covariant ones. Such a projection can be achieved by the G-twirling,

$$\mathcal{M} \to \bar{\mathcal{M}} = \int dg\mathcal{U}(g) \circ \mathcal{M} \circ \mathcal{U}^\dagger(g). \tag{3.26}$$

Similarly defining G-twirling for the states,

$$\rho \to \bar{\rho} = \int dg U(g)\rho U^\dagger(g), \tag{3.27}$$

so that $\bar{\rho}$ is a projection of ρ onto the set of U(1)-invariant (symmetric) states, one can rewrite the requirements 3.25a-3.25c as

$$\begin{aligned} \forall i \neq k : \bar{\mathcal{M}}_i(\rho_k) &= 0, & (3.28a) \\ \forall \rho : \sum_i \mathrm{Tr}[\bar{\mathcal{M}}_i(\rho)] &= 1. & (3.28b) \end{aligned}$$

The above conditions are however equivalent to

$$\begin{aligned} \forall i \neq k : \mathcal{M}_i(\bar{\rho}_k) &= 0, & (3.29a) \\ \forall \rho : \sum_i \mathrm{Tr}[\mathcal{M}_i(\bar{\rho})] &= 1, & (3.29b) \end{aligned}$$

which means that one needs to find unconstrained quantum operation $\mathcal{M}$ that allows for perfect unambiguous discrimination between G-twirled eigenstates of A.

This way the existence of a precise measurement of A is conditioned by the possibility of perfect distinguishibility between the set of states $\{\bar{\rho}_i\}$. One can discriminate $\{\bar{\rho}_i\}$

perfectly if and only if $\{\bar{\rho}_i\}$ have orthogonal supports. This is true only if all the states in G-twirled ensemble have rank one (as the number of states in the ensemble is equal to the dimension of the Hilbert space). Therefore one has to have

$$\forall i : \int dg U(g) |a_i\rangle \langle a_i| U^\dagger(g) = |\phi_i\rangle \langle \phi_i| . \tag{3.30}$$

Note however that pure states are the extremal points of the state space, i.e. they cannot be constructed as a convex combination of other states. Hence the only possibility to fulfill Eq. 3.30 is

$$\forall g \in G : |\phi_i\rangle \langle \phi_i| = |a_i\rangle \langle a_i| = U(g) |a_i\rangle \langle a_i| U^\dagger(g), \tag{3.31}$$

that is A must commute with N_1, which proofs the WAY theorem.

Chapter 4

Uncertainty Relations

In this chapter different results from the general field of uncertainty relations are presented and applied to the problems presented in the previous three chapters: relation between different asymmetric properties of a quantum state, the noise-disturbance relation and the WAY scenario measurement. First the oldest improvement of the Heisenberg uncertainty relation, known as the Schrödinger uncertainty relation, is described and applied to improve the bounds obtained by Ozawa. Then the notion of the skew information is introduced and linked with the asymmetry resources. This link allows to use the generalizations of the Heisenberg uncertainty relation for mixed states that use the skew information to obtain the bound for the asymmetry resources of a quantum state with respect to two non-commuting observables. These generalized relations are also used to improve Ozawa's bounds. Finally with the use of the entropic uncertainty relations another bound for asymmetric resources with respect to two different observables is obtained and the connection between the noise (disturbance) of quantum measurement and the mutually unbiased bases is presented.

4.1 Schrödinger-Robertson uncertainty relation

4.1.1 Derivation and justification

The standard derivation of the uncertainty relation is based on the use of the Schwartz inequality in the space of linear operators. Let's define the inner product of two operators A and B as

$$\langle A, B\rangle_\rho = \mathrm{Tr}\left(\rho A^\dagger B\right), \tag{4.1}$$

for which one can easily check the inner product defining properties: linearity, positive definiteness and conjugate symmetry. Now using the Schwartz inequality one obtains

$$\langle A, A\rangle_\rho \langle B, B\rangle_\rho \geq |\langle A, B\rangle_\rho|^2. \tag{4.2}$$

Defining the observables with subtracted average values,

$$\tilde{A} = A - \mathrm{Tr}\,(\rho A)\,\mathbb{1} = A - \langle A, \mathbb{1}\rangle_\rho \mathbb{1}, \tag{4.3a}$$

$$\tilde{B} = B - \mathrm{Tr}\,(\rho B)\,\mathbb{1} = B - \langle B, \mathbb{1}\rangle_\rho \mathbb{1}, \tag{4.3b}$$

one has

$$\langle \tilde{A}, \tilde{A}\rangle_\rho = \mathrm{Tr}\left(\rho A^2\right) - \left(\mathrm{Tr}\,(\rho A)\right)^2 = V(\rho, A), \tag{4.4a}$$

$$\langle \tilde{B}, \tilde{B}\rangle_\rho = \mathrm{Tr}\left(\rho B^2\right) - \left(\mathrm{Tr}\,(\rho B)\right)^2 = V(\rho, B), \tag{4.4b}$$

where $V(\rho, \cdot)$ denotes the variance in a state ρ, and

$$\langle \tilde{A}, \tilde{B}\rangle_\rho = \mathrm{Tr}\,(\rho AB) - \mathrm{Tr}\,(\rho A)\,\mathrm{Tr}\,(\rho B)\,. \tag{4.5}$$

Rewriting the product of A and B as the sum of commutator and anticommutator,

$$AB = \frac{1}{2}[A, B] + \frac{1}{2}\{A, B\}, \tag{4.6}$$

and using the Schwartz inequality, Eq. 4.2, for the operators $\tilde{A}$ and $\tilde{B}$ one arrives at

$$V(A)V(B) \geq \left|\frac{1}{2}\langle [A, B] + \{A, B\}\rangle - \langle A\rangle\langle B\rangle\right|^2, \tag{4.7}$$

where we dropped ρ assuming that all the averages are taken in the state ρ (unless stated otherwise) and $\langle\cdot\rangle$ denotes the average value. Using the fact that the expectation value of a commutator (anticommutator) is purely imaginary (real), the standard derivation of the uncertainty relation uses $|x|^2 = (\mathrm{Re}\,x)^2 + (\mathrm{Im}\,x)^2 \geq (\mathrm{Im}\,x)^2$ and thus gives the Heisenberg-Roberston uncertainty relation,

$$V(A)V(B) \geq \frac{1}{4}\left|\langle [A, B]\rangle\right|^2. \tag{4.8}$$

However, as was pointed out already in 1930 by Schrödinger [21], to fully exploit the Schwartz inequality one cannot omit the anticommutator part and should use

$$V(A)V(B) \geq \frac{\left|\langle [A, B]\rangle\right|^2 + \left|\langle\{A, B\}\rangle - 2\langle A\rangle\langle B\rangle\right|^2}{4}, \tag{4.9}$$

which is the Schrödinger-Robertson uncertainty relation.

The usual omission of the anticommutator part may come from the fact that for canonically conjugated variables x and p, i.e. $[x, p] = i\hbar$, the commutator part of the inequality never vanishes, whereas there exist states of minimum uncertainty for which the anticommutator part vanishes. However, as we are interested in the uncertainty relation for: (i) general observables (not only canonicaly conjugated ones), (ii) general states (not only the minimum uncertainty ones), we should not omit the anticommutator part and use the full uncertainty relation given by Eq. 4.9.

Too see the validity and the difference between the two uncertainty relations more clearly let us consider an example of a two-level system and two non-commuting observables given by Pauli matrices σ_z and σ_x. Let the (not necessarily pure) state of the system be described by the Bloch vector (n_x, n_y, n_z), so that $n_x^2 + n_y^2 + n_z^2 \leq 1$. In this simple case one can directly calculate the product of uncertainties as $\langle \sigma_\alpha^2 \rangle = 1$ and $\langle \sigma_\alpha \rangle^2 = n_\alpha^2$ (for $\alpha = x, z$), so that

$$V(\sigma_z)V(\sigma_x) = 1 - n_x^2 - n_z^2 + n_x^2 n_z^2 \geq n_y^2 + n_x^2 n_z^2, \tag{4.10}$$

and the equality holds only for pure states. Now one can compare the above result with the bounds given by the two uncertainty relations. The commutator of the considered observables is given by $[\sigma_z, \sigma_x] = 2\sigma_y$ and the anticommutator by $\{\sigma_z, \sigma_x\} = 0$. Therefore the Heisenberg uncertainty relation gives

$$V(\sigma_z)V(\sigma_x) \geq n_y^2, \tag{4.11}$$

whereas the Schrödinger uncertainty relation yields

$$V(\sigma_z)V(\sigma_x) \geq n_y^2 + n_z^2 n_x^2. \tag{4.12}$$

As can be seen the Schrödinger-Robertson relation yields a tight bound for all pure states, whereas the Heisenberg-Robertson relation fails to give any positive bound for the states in the XZ plane of the Bloch sphere. In general the bound given by Schrödinger uncertainty relation will always be tighter than the one given by the Heisenberg uncertainty relation.

Finally let us note that by defining the covariance of two observables A and B by

$$\mathrm{Cov}(A, B) = \left\langle \frac{AB + BA}{2} \right\rangle - \langle A \rangle \langle B \rangle, \tag{4.13}$$

which is the symmetrized version of covariance of two random variables (as the correct expression for the product of two non-commuting observables should be $(AB + BA)/2$), one can rewrite the Schrödinger-Robertson relation, Eq. 4.9, as

$$V(A)V(B) \geq \frac{1}{4} \left| \langle [A, B] \rangle \right|^2 + |\mathrm{Cov}(A, B)|^2, \tag{4.14}$$

or even more compactly by introducing the covariance matrix [21],

$$\boldsymbol{\sigma}(A, B) = \begin{pmatrix} V(A) & \mathrm{Cov}(A, B) \\ \mathrm{Cov}(A, B) & V(B) \end{pmatrix}, \tag{4.15}$$

to obtain

$$\det(\boldsymbol{\sigma}(A, B)) \geq \frac{1}{4} \left| \langle [A, B] \rangle \right|^2. \tag{4.16}$$

4.1.2 Application to Ozawa's bound for the WAY scenario

As described in Chapter 3, Wigner, Araki and Yanase showed that precise measurement of an observable A not commuting with the additive conserved quantity $L = L_1 \otimes \mathbb{1} + \mathbb{1} \otimes L_2$ is not possible [3, 4] and found the bound for the approximate measurement of such observable to be limited by the (appropriately defined) size of the measuring apparatus [18]. As Yanase used approximations in his derivation, the bound he found was not tight. Recently Ozawa in his paper [6] tried to find the tight limit of the measurement accuracy of the WAY-scenario observable A and claimed to obtain the "fundamental noise bound", which applied to the spin system gives "a correct lower bound, since our derivation uses no approximation". As we have described his derivation in Sec. 3.2, we just remind that the most important step of the derivation is based on the use of the Heisenberg-Robertson uncertainty relation to the noise operator N and the additive conserved quantity L, see Eq. 3.18. However we have just pointed out in the previous subsection that it is the Schrödinger-Robertson uncertainty relation (and not the Heisenberg-Robertson one) that gives a tight bound in a general case. If one replaces Heisenberg inequality by the improved Schödinger one and then follows Ozawa's derivation, instead of obtaining Eq. 3.21, one gets

$$\epsilon^2(|\psi\rangle) \geq \frac{\frac{1}{4}|\langle[A, L_1]\rangle|^2 + |\mathrm{Cov}(M, L_2) - \mathrm{Cov}(A, L_1)|^2}{V(L_1) + V(L_2)}. \tag{4.17}$$

In order to get the original Ozawa bound either both covariance terms must vanish or they must be equal. The first condition is not valid in general and the second one requires that the probe state $|\zeta\rangle$ must depend on the system state $|\psi\rangle$, which (i) is not possible if $|\psi\rangle$ is unknown (which is what Ozawa assumes); (ii) forbids the optimization of the probe state $|\zeta\rangle$ (the maximization of the variance is constrained by constant value of covariance).

It is interesting to note that the skew information that will be described in the next sections as a mean to improve the uncertainty relation for mixed states (i.e. to separate uncertainty coming from classical lack of knowledge due to mixture of states from the purely quantum uncertainty) was originally introduced by Wigner and Yanase [22] to quantify the "information which an ensemble described by a state vector or a statistical matrix contains with respect to the not easily measured quantities" (not easily measured quantity is defined by observable not commuting with the conserved quantity). Thus, the skew information may serve as a link between the WAY scenario considerations and the general uncertainty relations for mixed states.

4.1.3 Application to Ozawa's noise-disturbance uncertainty relation

In his derivation of the uncertainty relation for the noise and disturbance of the quantum measurement Ozawa again used the Heisenberg-Robertson inequality [2], which means that the error-disturbance bound can be tightened by improving the derivation with the use of the Schrödinger-Robertson relation. Unlike the WAY-scenario bound, however, here Ozawa did not aim to find the tight (fundamental) bound, but rather to get a universally valid and physically meaningful compact expression for the trade-off relation. Thus using the Heisenberg inequality instead of the Schrödinger one may have just been another approximation on the way to get the nice final expression. Still, it is interesting to check how the bound would be tightened if one used the Schrödinger uncertainty relation.

In order to see this, we will first briefly sketch Ozawa's original derivation. The idea is to use the definitions of noise and disturbance, given by Eqs. 2.23a-2.23b, and start from the commutator identity,

$$[N(A), D(B)] + [N(A), B] + [A, D(B)] = -[A, B]. \tag{4.18}$$

Then by taking the modulus of means of both sides and applying the triangular inequality twice one arrives at

$$|\langle[N(A), D(B)]\rangle| + |\langle[N(A), B]\rangle| + |\langle[A, D(B)]\rangle| \geq |\langle[A, B]\rangle|. \tag{4.19}$$

Now by applying Heisenberg-Robertson inequality to all of the terms on the left-hand-side of the above equation one gets

$$\sqrt{V(N(A))V(D(B))} + \sqrt{V(N(A))V(B)} + \sqrt{V(A)V(D(B))} \geq \frac{1}{2}|\langle[A, B]\rangle|. \tag{4.20}$$

Finally by noting that $\epsilon(A) = \langle N^2(A)\rangle^{\frac{1}{2}} \geq \sqrt{V(N(A))}$ and $\eta(B) = \langle D^2(B)\rangle^{\frac{1}{2}} \geq \sqrt{V(D(B))}$ (i.e. the average square is bigger than variance) one obtains

$$\epsilon(A)\eta(B) + \epsilon(A)\sqrt{V(B)} + \sqrt{V(A)}\eta(B) \geq \frac{1}{2}|\langle[A, B]\rangle|, \tag{4.21}$$

which is exactly Ozawa's noise-disturbance relation presented in Sec. 2.4.1 and given by Eq. 2.26 (only with slightly different notation, $\sigma \to \sqrt{V}$).

The improvement must be done at the level of Eq. 4.19 by applying the Schrödinger-Robertson inequality (its most compact version, Eq. 4.16) to all of the terms on the left hand side,

$$[\det(\boldsymbol{\sigma}(N(A), D(B)))]^{\frac{1}{2}} + [\det(\boldsymbol{\sigma}(N(A), B))]^{\frac{1}{2}} + [\det(\boldsymbol{\sigma}(A, D(B)))]^{\frac{1}{2}} \geq \frac{1}{2}|\langle[A, B]\rangle|. \tag{4.22}$$

4.2 Skew information

Skew information was first introduced by Wigner and Yanase [22] as a non-entropic based measure of information that a given state contains with respect to observables not commuting with the conserved quantity. They were motivated by the WAY theorem and the fact that such observables are "not easily measurable", so that the entropy, which is a measure of our total knowledge into which the knowledge of the value of any observable enters in the same way, is not the best way to describe information content. They postulated the requirements that an expression for the information content I should satisfy:

1. $I(p\rho_1 + (1-p)\rho_2) \leq pI(\rho_1) + (1-p)I(\rho_2)$,
2. $I(\rho) \geq I(\rho_1) + I(\rho_2)$ for $\rho_1 = \mathrm{Tr}_2\rho$ and $\rho_2 = \mathrm{Tr}_1\rho$,
3. The information content of the isolated system should be constant in time.

The first point, the convexity of function I, is physically motivated by the fact that the information content of the union of two different ensembles ρ_1 and ρ_2 should be less than average information content of the component ensembles, because by uniting the two systems one "forgets" from which ensemble a particular sample comes from. The second point is motivated by the fact that by separating two subsystems one "forgets" about any correllations, so that the equality is obtained only for $\rho = \rho_1 \otimes \rho_2$. The last requirement comes from the fact that both quantum and classical evolution of the closed system is described by the causal equations, so that the information about the isolated system at a given time yields this information at any other time.

As a measure of information content satisfying the above conditions for pure state Wigner and Yanase proposed the following definition

$$I(|\psi\rangle\langle\psi|, A) = \mathrm{Tr}\left(|\psi\rangle\langle\psi| A^2\right) - (\mathrm{Tr}(|\psi\rangle\langle\psi| A))^2, \tag{4.23}$$

which is exactly the same as the variance of A in state $|\psi\rangle$, $I(A) = V(A)$. They argued that in the presence of a conservation law the information content of all pure states should not be the same, which is the case when one is using the von Neumann entropy. Although they wrote: "whereas a characteristic vector of the conserved quantity contains no such (skew) information, a state vector which lies skew to these characteristic vectors does", they did not give any operative meaning for the skew information. However, as we showed in Sec. 1.2, if one restricts to pure states, then the variance (which is equivalent to the skew information) of a state is an asymmetry monotone, and as such may be used to quantify the asymmetry resource (with respect to the symmetry generated by operator of a conserved quantity), so any of the operative meanings of such resource may be used. For example $I(|\psi\rangle, A)$, the skew information of state $|\psi\rangle$ with respect to observable A, may quantify the information that can be encoded by this state using symmetric quantum operations or the information loss after the G-twirling of $|\psi\rangle$.

There are many ways in which one may extend the definition of skew information, Eq. 4.23, to the mixed states, so that it still fulfills the requirements for the measure of information content. The original extended definition [22] is given by

$$I(\rho, A) = -\frac{1}{2}\mathrm{Tr}\left([\sqrt{\rho}, A]^2\right) = \mathrm{Tr}\left(\rho A^2\right) - \mathrm{Tr}\left(\sqrt{\rho}A\sqrt{\rho}A\right), \tag{4.24}$$

and is known as Wigner-Yanase skew information. The more general version of the above expression is known as Wigner-Yanase-Dyson skew information and is given by

$$I_s(\rho, A) = -\frac{1}{2}\mathrm{Tr}\left([\rho^s, A][\rho^{1-s}, A]\right) = \mathrm{Tr}\left(\rho A^2\right) - \mathrm{Tr}\left(\rho^s A \rho^{1-s} A\right). \tag{4.25}$$

It was proven by Lieb [23] that the above expression is convex in ρ, therefore it fulfills the basic requirement for the information content (the other two postulated requirements are also satisfied).

The reason why this thesis contains a section about the skew information is because both Wigner-Yanase and Wigner-Yanase-Dyson skew information are the asymmetry monotones and can be used to quantify asymmetry resource with respect to a given U(1) symmetry. The proof is based on the construction of an asymmetry monotone from the information monotone, in this case the relative Renyi entropy, and is neatly presented in Ref. [7]. Because of that one can use the results concerning the skew information to get insight into the asymmetric properties of quantum states. One of the fields where the skew information has recently been used is the field of the Heisenberg-like uncertainty relations for mixed states, which will be described in the following section.

4.3 Uncertainty relations for mixed states

4.3.1 Motivation and separation of uncertainties

Recently new uncertainty measures were used to enhance the Heisenberg and Schrödinger uncertainty relations for mixed states. The main motivation is that the total uncertainty of an observable A in a mixed state ρ, V_{tot}, comes both from classical mixing C and quantum uncertainty Q, where of course for any measure used one should have $V_{tot} > C$ and $V_{tot} > Q$. The goal now is to find Heisenberg-like inequality for Q (or C), which (due to $Q, C < V_{tot}$) will be a tighter bound than the original uncertainty relation. Luo [24] proposed to separate the standard uncertainty measure, the variance $V(\rho, A)$, into classical, $C(\rho, A)$, and quantum, $Q(\rho, A)$, compontents in a linear fashion,

$$V(\rho, A) = C(\rho, A) + Q(\rho, A). \tag{4.26}$$

He also suggested that C and Q should satisfy the following requirements:

1. $Q(p\rho_1 + (1-p)\rho_2, A) \leq pQ(\rho_1, A) + (1-p)Q(\rho_2, A)$,

2. $C(p\rho_1 + (1-p)\rho_2, A) \geq pC(\rho_1, A) + (1-p)C(\rho_2, A)$,

3. $Q(\rho, A) = V(\rho, A)$ and $C(\rho, A) = 0$ for $\rho = |\psi\rangle \langle\psi|$,

4. $C(\rho, A) = V(\rho, A)$ and $Q(\rho, A) = 0$ for $[A, \rho] = 0$.

The justification of the first two requirements is that classical mixing should (should not) increase classical (quantum) uncertainty. The third point is justified, because there is no classical mixing in a pure state, so the whole uncertainty must be quantum. The last requirement comes from the fact that when ρ and A commute they can be simultaneously diagonalized and any eigenstate of ρ is an eigenstate of A, therefore, since ρ is a classical mixture of its eigenstates, the quantum uncertainty should vanish.

Let us note that the separation of the total uncertainty into a sum of the classical and quantum one (and not into some other function of C and Q) not only has a clear physical meaning, but also has the advantage that by defining quantum uncertainty Q that satisfies the aforementioned requirements, the classical uncertainty C fulfills these requirements automatically. To see these let us first show that the variance is a concave function of ρ, because

$$\begin{aligned}
V(p\rho_1 + (1-p)\rho_2, A) &= p\langle A^2\rangle_{\rho_1} + (1-p)\langle A^2\rangle_{\rho_2} - (p\langle A\rangle_{\rho_1} + (1-p)\langle A\rangle_{\rho_2})^2, \\
pV(\rho_1, A) + (1-p)V(\rho_2, A) &= p(\langle A^2\rangle_{\rho_1} - \langle A\rangle^2_{\rho_1}) + (1-p)(\langle A^2\rangle_{\rho_2} - \langle A\rangle^2_{\rho_2}),
\end{aligned}$$

and by subtracting the second of the above equations from the first we get

$$\begin{aligned}
&p\langle A\rangle^2_{\rho_1} + (1-p)\langle A\rangle^2_{\rho_2} - p^2\langle A\rangle^2_{\rho_1} - (1-p)^2\langle A\rangle^2_{\rho_2} - 2p(1-p)\langle A\rangle_{\rho_1}\langle A\rangle_{\rho_2} = \\
&p(1-p)\langle A\rangle^2_{\rho_1} + p(1-p)\langle A\rangle^2_{\rho_2} - 2p(1-p)\langle A\rangle_{\rho_1}\langle A\rangle_{\rho_2} = p(1-p)(\langle A\rangle_{\rho_1} - \langle A\rangle_{\rho_2})^2 \geq 0,
\end{aligned}$$

so that

$$V(p\rho_1 + (1-p)\rho_2, A) \geq pV(\rho_1, A) + (1-p)V(\rho_2, A). \tag{4.27}$$

Now, if Q is convex then from concavity of the variance, the concavity of C follows straightforwardly. The third and fourth requirements are also automatically fulfilled for C if Q satisfies them.

4.3.2 Skew information based uncertainty relations

The separation of uncertainties proposed by Luo leaves us only with the ambiguity of choosing the measure of the quantum uncertainty, Q, satisfying the aforementioned conditions, as the classical uncertainty will be then defined by $C = V - Q$. As the skew information satisfies all requirements for the quantum uncertainty, it defines the following separation:

$$Q(\rho, A) = \mathrm{Tr}\left(\rho A^2\right) - \mathrm{Tr}\left(\sqrt{\rho} A \sqrt{\rho} A\right), \tag{4.28a}$$

$$C(\rho, A) = \mathrm{Tr}\left(\sqrt{\rho} A_0 \sqrt{\rho} A_0\right). \tag{4.28b}$$

To find the improved uncertainty relation with the quantities defined above let us mix the derivations presented in Refs. [25] and [26], so that one can easily obtain the two results presented in these papers. We first define

$$A_{\rho+} = \frac{1}{\sqrt{2}}\{\sqrt{\rho}, A_0\} \tag{4.29a}$$

$$A_{\rho-} = \frac{1}{\sqrt{2}}[\sqrt{\rho}, A_0] \tag{4.29b}$$

Now using the Hilbert-Schmidt inner product,

$$\langle A, B\rangle = \mathrm{Tr}\left(A^\dagger B\right), \tag{4.30}$$

one has

$$||A_{\rho+}||^2 = \frac{1}{2}\mathrm{Tr}\left(\{\sqrt{\rho}, A_0\}^2\right) = J(\rho, A_0), \tag{4.31a}$$

$$||A_{\rho-}||^2 = -\frac{1}{2}\mathrm{Tr}\left([\sqrt{\rho}, A_0]^2\right) = I(\rho, A_0), \tag{4.31b}$$

$$\langle A_{\rho+} A_{\rho-}\rangle = 0, \tag{4.31c}$$

where $I(\rho, A_0)$ is Wigner-Yanase skew information and $J(\rho, A_0)$ is a quantity conjugated to it. Let us also note that

$$I(\rho, A_0) = I(\rho, A), \tag{4.32a}$$

$$J(\rho, A_0) = J(\rho, A) - 2(\mathrm{Tr}\,(\rho A))^2. \tag{4.32b}$$

Now using the Schwartz inequalities for operators $A_{\rho+}$, $B_{\rho-}$ and $A_{\rho-}$, $B_{\rho+}$ one can obtain the following relations:

$$I(\rho, A) J(\rho, B_0) \geq \frac{1}{4}|\langle [A, B]\rangle|^2, \tag{4.33a}$$

$$I(\rho, B) J(\rho, A_0) \geq \frac{1}{4}|\langle [A, B]\rangle|^2. \tag{4.33b}$$

By adding or multiplying both sides of the above inequalities one can obtain the result of Ref. [25] and [26], respectively. The addition gives

$$V(\rho, A) V(\rho, B) \geq \frac{1}{4}|\langle [A, B]\rangle|^2 + C(\rho, A) C(\rho, B), \tag{4.34}$$

and the multiplication yields

$$U(\rho, A) U(\rho, B) \geq \frac{1}{4}|\langle [A, B]\rangle|^2, \tag{4.35}$$

where

$$U(\rho, A) = \sqrt{V^2(\rho, A) - (V(\rho, A) - I(\rho, A))^2}. \tag{4.36}$$

The first of the above uncertainty relations is the Heisenberg-Robertson inequality, Eq. 4.8, with the bound improved by the addition of the term coming from classical uncertainty. It is important to note that Park in his paper [26] also found the similarly improved (i.e. by adding new terms to the bound) Schrödinger-Robertson inequality, however the additional terms are far more complicated than $C(\rho, A)C(\rho, B)$ and required approximations.

The second uncertainty relation is defined in terms of the new measure of uncertainty, $U(\rho, A)$, which may be treated as quantum uncertainty, because $U = [V^2 - (V - C)^2]^{\frac{1}{2}}$, so it measures the difference between total and classical uncertainty (which quantum uncertainty should be). It can also be seen as another way of separating the uncertainty into the classical and quantum components in a quadratic fashion,

$$V^2(\rho, A) = C^2(\rho, A) + U^2(\rho, A). \tag{4.37}$$

Therefore Eq. 4.35 is also an improvement of the Heisenberg-Robertson inequality (as $U \leq V$), with the interpretation that the standard commutator bound bounds not only the total, but also the quantum uncertainty.

Both of these uncertainty relations may be used to improve Ozawa's universally valid error-uncertainty relations for mixed states. Especially the second one, because of the same form as the Heisenberg-Robertson uncertainty relation used by Ozawa in his derivation, leads to a nice result for the trade-off relation,

$$\epsilon(A)\eta(B) + \epsilon(A)U(B) + U(A)\eta(B) \geq \frac{1}{2}|\langle[A, B]\rangle|, \tag{4.38}$$

which is exactly the original relation, Eq. 2.26, with the total uncertainty V exchanged by the quantum uncertainty U. Moreover Furuichi [27] recently derived the Schrödinger-Robertson-like relation in terms of U,

$$U(\rho, A)U(\rho, B) \geq \frac{1}{4}|\langle[A, B]\rangle|^2 + |\mathrm{Corr}_\rho(A, B)|^2, \tag{4.39}$$

where

$$\mathrm{Corr}_\rho(A, B) = \mathrm{Tr}\left(\rho\frac{\{A, B\}}{2}\right) - \mathrm{Tr}\left(\sqrt{\rho}A\sqrt{\rho}B\right) = \mathrm{Tr}\left(\rho\frac{\{A_0, B_0\}}{2}\right) - \mathrm{Tr}\left(\sqrt{\rho}A_0\sqrt{\rho}B_0\right) \tag{4.40}$$

is the correlation term that is the counterpart of the original covariance term appearing in Eq. 4.14. To understand the meaning of this additional term let us follow the reasoning presented in Ref. [24]. Similarly as the variance V (measure of total uncertainty) can be split into the skew information I (measure of quantum uncertainty) and C (measure of classical uncertainty), the covariance can also be split in the corresponding way (see Eqs. 4.28a-4.28b),

$$\begin{aligned} \mathrm{Cov}_\rho(A,B) &= \mathcal{Q}_\rho(A,B) + \mathcal{C}_\rho(A,B), && (4.41a)\\ \mathcal{Q}_\rho(A,B) &= \mathrm{Tr}\left(\rho\frac{\{A_0,B_0\}}{2}\right) - \mathrm{Tr}\left(\sqrt{\rho}A_0\sqrt{\rho}B_0\right), && (4.41b)\\ \mathcal{C}_\rho(A,B) &= \mathrm{Tr}\left(\sqrt{\rho}A_0\sqrt{\rho}B_0\right). && (4.41c) \end{aligned}$$

Therefore the correlation term appearing in Eq. 4.39 can be interpreted as the quantum part of the covariance (anticommutator) term appearing in the original Schrödinger-Robertson inequality. Hence one can improve the Ozawa's bound even further, by using both improvements: use the Schrödinger-Robertson relation instead of Heisenberg-Robertson one and use the quantum component of uncertainties and covariances instead of the total ones.

Note that natural improvement of the uncertainty relation with the use of the skew information alone,

$$I(\rho,A)I(\rho,B) \geq \frac{1}{4}|\langle[A,B]\rangle|^2, \tag{4.42}$$

is not valid. However using the Schwartz inequality for operators A_{ρ_-} and B_{ρ_-} (see Eq. 4.29b) one finds

$$I(\rho,A)I(\rho,B) \geq \mathcal{Q}_\rho(A,B), \tag{4.43}$$

so that the product of skew informations with respect to different observables is bounded by the quantum component of the covariance of these observables. As the skew information is the asymmetry monotone, the above result establishes a relation between asymmetric properties of a quantum state with respect to two non-commuting observables. For better understanding of the link between uncertainties and the asymmetry it is vital to check whether $U(\rho,A)$ is also an asymmetry monotone with respect to the symmetry generated by A. If it is the case then, since U "obeys" the Heisenberg-Robertson uncertainty relation used by Ozawa, both Ozawa's results (the error-disturbance relation and the precision of the WAY scenario measurement) can be redefined in terms of asymmetry monotones.

One more uncertainty relation that uses the skew information can be easily derived [7],

$$I(\rho,A)V(\rho,B) \geq \frac{1}{8}|\langle[A,B]\rangle_\rho|^2. \tag{4.44}$$

This is the least refined inequality from the ones presented in this section, however it can be straightforwardly applied to change the Ozawa bound for the WAY scenario. Instead of Eq. 3.21 one gets

$$\epsilon^2(\rho) \geq \frac{|\langle[A,L_1]\rangle|^2}{8I(\rho,L_1)+8I(\rho_{probe},L_2)}, \tag{4.45}$$

which is a stronger bound for mixed states for which $V > 2I$.

Some of the results described in this section can be further extended by generalizing the Wigner-Yanase skew information to Wigner-Yanase-Dyson (WYD) skew information [28]. The WYD skew information, I_s, and the quantity conjugated to it, J_s (similarly to Eq. 4.31a), are defined by

$$I_s(\rho, A) = \mathrm{Tr}\left(\rho A_0^2\right) - \mathrm{Tr}\left(\rho^s A_0 \rho^{1-s} A_0\right), \tag{4.46a}$$

$$J_s(\rho, A) = \mathrm{Tr}\left(\rho A_0^2\right) + \mathrm{Tr}\left(\rho^s A_0 \rho^{1-s} A_0\right). \tag{4.46b}$$

Introducing $l_s(\rho, A, B)$,

$$l_s(\rho, A, B) = \mathrm{Tr}\left(\rho[A, B]\right) - \mathrm{Tr}\left(\rho^{|2s-1|}[A, B]\right), \tag{4.47}$$

one can prove that

$$I_s(\rho, A)J_s(\rho, B) \geq \frac{1}{4}|l_s(\rho, A, B)|^2, \tag{4.48a}$$

$$I_s(\rho, B)J_s(\rho, A) \geq \frac{1}{4}|l_s(\rho, A, B)|^2, \tag{4.48b}$$

so that also

$$U_s(\rho, A)U_s(\rho, B) \geq \frac{1}{4}|l_s(\rho, A, B)|^2, \tag{4.49}$$

where

$$U_s(\rho, A) = \sqrt{V^2(\rho, A) - (V(\rho, A) - I_s(\rho, A))^2}. \tag{4.50}$$

Thus the standard commutator bound is replaced by $l_s(\rho, A, B)$.

4.4 Entropic uncertainty relations

In 1983 Deutsch [29] proposed and motivated an alternative approach to the uncertainty relations that uses the Shannon entropy to quantify uncertainty. Later, in 1988, Maassen and Uffink [30] proved the generalized entropic uncertainty relation of the following form

$$\frac{1}{2}(H(A, \rho) + H(B, \rho)) \geq -\log c(A, B), \tag{4.51}$$

where $H(A, \rho)$ is the Shannon entropy of the probability distributon of the measurement outcomes of observable A in a state ρ and

$$c(A, B) = \max_{n,m}\{|\langle a_n|b_m\rangle| : A\,|a_n\rangle = a_n\,|a_n\rangle\,,\; B\,|b_n\rangle = b_n\,|b_n\rangle\,, \tag{4.52}$$

so it is the maximum overlap between the eigenstates of observables A and B. Comparing the definition of $H(A, \rho)$ with the expression for the relative entropy of frameness for pure states, Eq. 1.11, one can find out that these are the same. Therefore one can use the above entropic uncertainty relation to establish another relation between asymmetric resources of a quantum state with respect to two non-commuting observables,

$$As_A(|\psi\rangle\langle\psi|) + As_B(|\psi\rangle\langle\psi|) \geq -2\log c(A,B). \tag{4.53}$$

It is worth noting that the strongest bound [the largest $c(A,B)$] is obtained for the observables, eigenstates of which form the mutually unbiased bases, i.e.

$$\forall_n\, |a_n\rangle = \frac{1}{\sqrt{d}}\sum_m e^{i\phi_{mn}}\,|b_m\rangle\,, \tag{4.54}$$

where d is the dimension of the Hilbert space. The maximal bound is therefore equal to $\log d$. The classical contexts connected with such observables may be called mutually excluding contexts, as having a state with definite value of one observable means that the value of the other one is maximally uncertain (uniformly distributed). A general theorem states that for every finite dimensional Hilbert space there exists at least three mutually unbiased bases [31]. This means that one can always find non-commuting mutually unbiased observables A and B and a state $|\psi\rangle$ from the third unbiased basis, so that

$$\forall_n |\langle a_n|\psi\rangle|^2 = |\langle b_n|\psi\rangle|^2 = \frac{1}{d}. \tag{4.55}$$

There are two interesting facts connected with this result. First, since the probability distribution of the outcomes are the same for A and B, therefore, according to the notion of a state-dependent measurement noise introduced in Sec. 2.1, if instead of measuring A one measures B, this measurement will be treated as noiseless. Similarly the measurement of A will not affect the probability distribution of the measurement of B (with and without preceeding measurement of A the probability distribution of the outcomes is uniform), so according to the notion of a state-dependent measurement disturbance introduced in Sec. 2.1, the measurement of B is not disturbed by the measurement of A. This means that for every finite-dimensional Hilbert space there exist a state for which two non-commuting observables can be measured without any noise or disturbance. The second interesting fact is that for two mutually unbiased observables A and B there always exists a state $|\psi\rangle$ that maximizes the asymmetry resources with respect to U(1) symmetries generated by both observables, i.e.

$$As_A(|\psi\rangle\langle\psi|) = As_B(|\psi\rangle\langle\psi|) = \log d. \tag{4.56}$$

Bibliography

[1] H. P. Robertson, Physical Review **34**, 163 (1929).

[2] M. Ozawa, Physical Review A **67**, 042105 (2003).

[3] P. Busch, arXiv preprint arXiv:1012.4372 (2010).

[4] H. Araki and M. M. Yanase, Physical Review **120**, 622 (1960).

[5] E. Noether, Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen 235 (1918).

[6] M. Ozawa, Physical Review Letters **88**, 050402 (2002).

[7] I. Marvian, (2012).

[8] I. Marvian and R. W. Spekkens, New Journal of Physics **15**, 033001 (2013).

[9] P. Busch, P. Lahti, and R. F. Werner, arXiv preprint arXiv:1306.1565 (2013).

[10] M. Ahmadi, D. Jennings, and T. Rudolph, New Journal of Physics **15**, 013057 (2013).

[11] I. Marvian and R. W. Spekkens, arXiv preprint arXiv:1212.3378 (2012).

[12] G. Gour, I. Marvian, and R. W. Spekkens, Physical Review A **80**, 012307 (2009).

[13] G. Gour and R. W. Spekkens, New Journal of Physics **10**, 033023 (2008).

[14] W. K. Wootters and W. H. Zurek, Nature **299**, 802 (1982).

[15] M. Ozawa, Journal of Optics B: Quantum and Semiclassical Optics **7**, S672 (2005).

[16] G. Sulyok, S. Sponar, J. Erhart, G. Badurek, M. Ozawa, and Y. Hasegawa, arXiv preprint arXiv:1305.7251 (2013).

[17] L. A. Rozema, D. H. Mahler, A. Hayat, and A. M. Steinberg, arXiv preprint arXiv:1307.3604 (2013).

[18] M. M. Yanase, Physical Review **123**, 666 (1961).

[19] L. Loveridge and P. Busch, The European Physical Journal D **62**, 297 (2011).

[20] P. Busch and L. Loveridge, Physical Review Letters **106**, 110406 (2011).

[21] E. Schrödinger, arXiv preprint quant-ph/9903100 (1999).

[22] E. P. Wigner and M. M. Yanase, Proceedings of the National Academy of Sciences of the United States of America **49**, 910 (1963).

[23] E. H. Lieb, Advances in Mathematics **11**, 267 (1973).

[24] S. Luo, Theoretical and mathematical physics **143**, 681 (2005).

[25] S. Luo, Physical Review A **72**, 042110 (2005).

[26] Y. M. Park, Journal of mathematical physics **46**, 042109 (2005).

[27] S. Furuichi, arXiv preprint arXiv:1005.2655 (2010).

[28] D. Li, X. Li, F. Wang, H. Huang, and L. Kwek, Physical Review A **79**, 052106 (2009).

[29] D. Deutsch, Physical Review Letters **50**, 631 (1983).

[30] H. Maassen and J. B. Uffink, Physical Review Letters **60**, 1103 (1988).

[31] S. Bandyopadhyay, P. O. Boykin, V. Roychowdhury, and F. Vatan, Algorithmica **34**, 512 (2002).