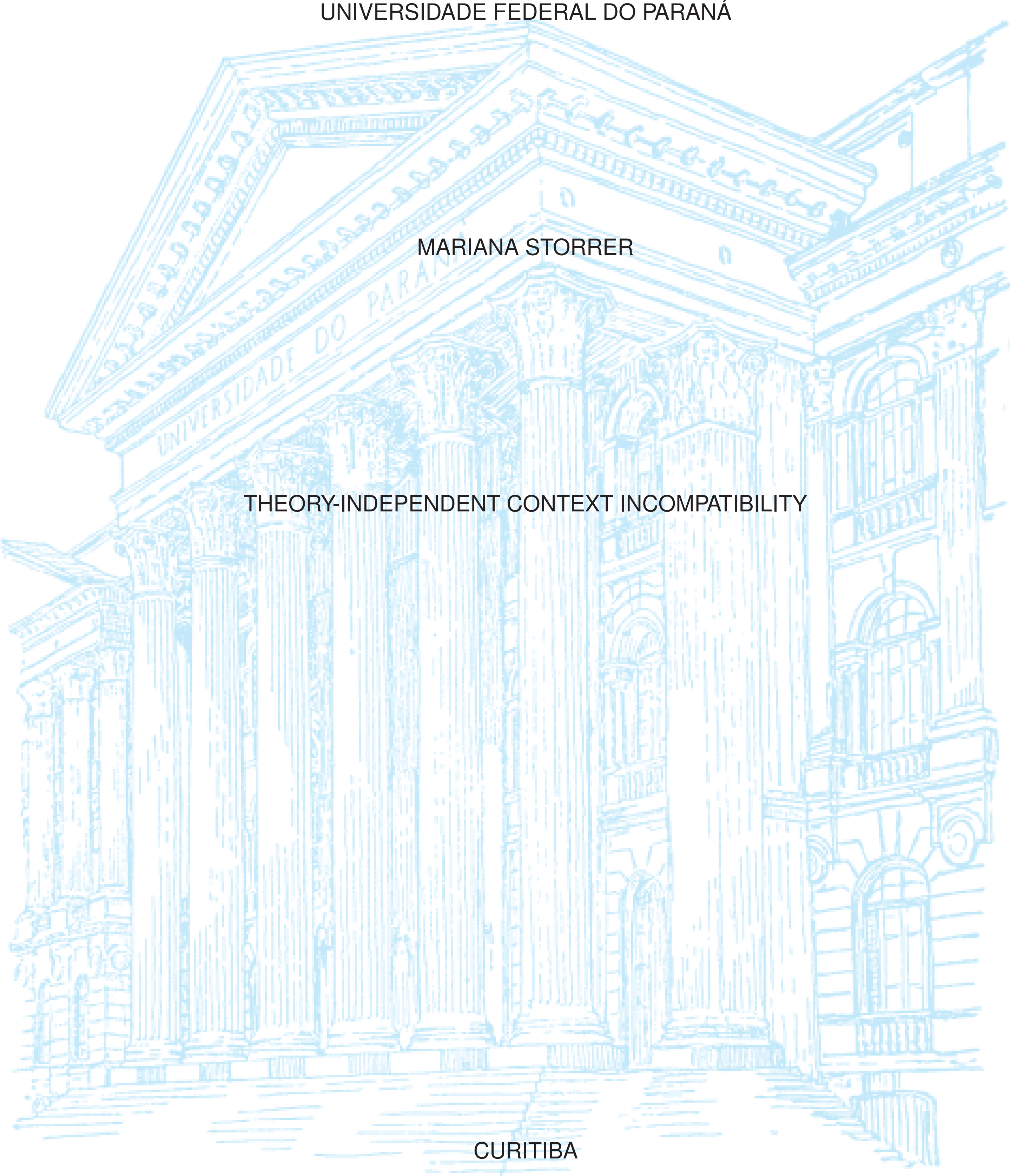
UNIVERSIDADE FEDERAL DO PARANÁ



MARIANA STORRER

THEORY-INDEPENDENT CONTEXT INCOMPATIBILITY

CURITIBA

2023

MARIANA STORRER

THEORY-INDEPENDENT CONTEXT INCOMPATIBILITY

Dissertação apresentada como requisito parcial à obtenção do título de Mestre ao Programa de Pós-Graduação em Física do Setor de Ciências Exatas na Universidade Federal do Paraná.

Orientador: Prof. Dr. Renato Moreira Ângelo

CURITIBA

2023

DADOS INTERNACIONAIS DE CATALOGAÇÃO NA PUBLICAÇÃO (CIP)
UNIVERSIDADE FEDERAL DO PARANÁ
SISTEMA DE BIBLIOTECAS – BIBLIOTECA CIÊNCIA E TECNOLOGIA

Storrer, Mariana.
Theory-independent context incompatibility. / Mariana Storrer. – Curitiba, 2023.
1 recurso on-line : PDF.

Dissertação (Mestrado) – Universidade Federal do Paraná, Setor de Ciências Exatas, Programa de Pós-Graduação em Física.
Orientador: Prof. Dr. Renato Moreira Ângelo.

1. Física. 2. Teoria quântica. 3. Mecânica quântica. 4. Entropia quântica. I. Ângelo, Renato Moreira. II. Universidade Federal do Paraná. Programa de Pós-Graduação em Física. III. Título.

Bibliotecário: Nilson Carlos Vieira Júnior CRB-9/1797

MINISTÉRIO DA EDUCAÇÃO
SETOR DE CIENCIAS EXATAS
UNIVERSIDADE FEDERAL DO PARANÁ
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO FÍSICA - 40001016020P4

TERMO DE APROVAÇÃO

Os membros da Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação FÍSICA da Universidade Federal do Paraná foram convocados para realizar a arguição da dissertação de Mestrado de **MARIANA STORRER** intitulada: **"Theory-independent context incompatibility"**, sob orientação do Prof. Dr. RENATO MOREIRA ANGELO, que após terem inquirido a aluna e realizada a avaliação do trabalho, são de parecer pela sua APROVAÇÃO no rito de defesa.
A outorga do título de mestra está sujeita à homologação pelo colegiado, ao atendimento de todas as indicações e correções solicitadas pela banca e ao pleno atendimento das demandas regimentais do Programa de Pós-Graduação.

CURITIBA, 24 de Fevereiro de 2023.

Assinatura Eletrônica
25/02/2023 08:40:32.0
RENATO MOREIRA ANGELO
Presidente da Banca Examinadora

Assinatura Eletrônica
24/02/2023 17:54:04.0
GIOVANI LOPES VASCONCELOS
Avaliador Interno (UNIVERSIDADE FEDERAL DO PARANÁ)

Assinatura Eletrônica
27/02/2023 18:01:43.0
FERNANDO ROBERTO DE LUNA PARISIO FILHO
Avaliador Externo (UNIVERSIDADE FEDERAL DE PERNAMBUCO)

ACKNOWLEDGEMENTS

To my family, first, as they have always supported me in more ways than I could count, making this path a possibility for me. A special thank you to my sister who read some of this work to make sure my writing was understandable.

To my boyfriend, who encouraged me through everything.

To Prof. Renato Angelo and Prof. Ana Costa, for all of their work, contribution, and enormous patience in trying to teach me how to make proper research.

To the group of Foundations of Quantum Mechanics at UFPR, especially Henrique, for the discussions and input.

To anyone that read all of this (yes, you).

Finally, to Capes, for the funding, and to the graduate program in Physics.

RESUMO

Acredita-se que incompatibilidade esteja no cerne de “estranhezas” quânticas, permitindo, junto com emaranhamento, efeitos quânticos que desafiam nossa percepção da natureza, como as violações de desigualdades de Bell. Não obstante sua importância, não há uma única definição e muitas abordagens foram propostas. Ainda sim, algo se destaca nelas: a dependência na relação entre as medições, sem considerar o estado que está sendo medido. Como estados essencialmente clássicos quase não são perturbados por medições, em contraste com o que ocorre para estados quânticos, parece negligente desconsiderá-los na definição de incompatibilidade. Assim, recentemente foi proposta a *Incompatibilidade de um Contexto Físico*, onde o contexto nada mais é que o conjunto formado pelas medições a serem realizadas e pelo estado, e é o conjunto todo que é dito incompatível ou não. Esta nova abordagem, no entanto, foi feita em uma estrutura estritamente quântica, o que limita seu uso em possíveis futuras teorias pós-quânticas que podem criar correlações ainda mais fortes que as quânticas. Nós propomos neste trabalho uma definição para incompatibilidade de contexto que é inteiramente independente de teoria, baseada apenas em probabilidades e mapas de medição não seletiva. A definição propõe um par de equações que qualquer contexto compatível deve satisfazer, onde mostramos então um regime clássico que as satisfazem prontamente e um estudo de caso para o qubit na representação de Bloch para ilustrar quando contextos quânticos as satisfazem e, consequentemente, quando as violam. Nós também construimos quantificadores: um baseado na entropia relativa de von Neumann e um intercambiável baseado na divergência de Kullback-Leibler para probabilidades. Eles nos permitem comparar o que a definição proposta quantifica contra o que a definição já existente quantificava, validando nossa abordagem como comparável mas não igual a ela. Também mostramos que incompatibilidade de medição está contida na nossa definição como uma escolha específica de contexto, o que solidifica a proposta apresentada nessa dissertação.

Key-words: mecânica quântica; incompatibilidade; incompatibilidade de contexto; incompatibilidade de contexto independente de teoria.

ABSTRACT

Incompatibility is believed to be at the center of quantum “weirdness”. It has been shown to enable, together with entanglement, quantum effects that challenge our perception of Nature such as the Bell inequality violation. Notwithstanding its importance, incompatibility is not unanimously defined, and many different approaches have been conceptualized. Nevertheless, one thing stands out: many of these approaches rely only on the role of the measurements, without considering the preparation. Since classically behaving states are nearly not disturbed by measurements in the way quantum behaving states are, it seems neglectful to disregard the states in the definition of incompatibility. Hence the recently proposed *Incompatibility of a Physical Context*, with the context being the set of measurements and the state of the system, where now it is a context that is incompatible or not. This approach, however, is constrained to the quantum framework, which limits its use in forthcoming notions of post-quantum theories that can elevate quantum correlations to stronger ones. We propose, then, a new definition for context incompatibility that is entirely independent on the theory, based on probabilities and non-selective measurement maps only. The definition poses a set of equations that any compatible context must obey. We then show a classical regime that readily satisfies them and a case study for the qubit in the Bloch representation to illustrate when a quantum context would satisfy them and, consequently, when it would not. We also built quantifiers: one based on the von Neumann relative entropy and an interchangeable one based on the Kullback-Leibler divergence for probabilities. They allow us to compare what our definition quantifies against what the existing context incompatibility definition quantified, validating our approach as comparable but not equal. We also show that incompatibility of measurements is contained in our proposed definition as a specific choice of context, which further solidifies the proposal presented in this dissertation.

Palavras-chaves: quantum mechanics; theory-independent context incompatibility; incompatibility; context incompatibility.

LIST OF FIGURES

CONTENTS

INTRODUCTION

Consider a ball in a football game, like the one the Brazilian team lost and was thus eliminated from the World Cup. A viewer of such a game could always tell where the ball was, and know roughly its speed – sometimes faster, sometimes slower. That is, in fact, a very important part of the game. Now, imagine that whenever the spectator focused on seeing where exactly the ball was in the field, they could not tell you its speed, or if it was in movement at all. This is what would happen if position and momentum were *incompatible* in the classical realm as they are in the quantum realm. Brazil could have been saved if only that were a quantum game.

But, rather, incompatibility is this odd effect that occurs only in quantum systems and, as one can tell from this example, can be very counter-intuitive. It has been noticed in quantum theory at least as early as Heisenberg's work in 1927 [1] on the uncertainty relation $\Delta x \Delta p \gtrsim h$, where he realized that the wave packets describing the position and the momentum cannot be arbitrarily small, having a limitation of nearly Planck's constant h. Incompatibility appears in some form in many seminal works of quantum mechanics, such as Bohr's complementary principle [2], when he speaks of the measurement disturbing the system; EPR's paradox [3], where the authors explicitly use position and momentum and its incompatible quantum nature alongside the entanglement property of some quantum states to infer that quantum theory was incomplete; and the violation of Bell inequalities can only be achieved through incompatible measurements [4–6], which were used in the ultimate loop-hole free experimental test [7], cementing that quantum mechanics does violate a Bell-type inequality and thus cannot be thoroughly described by a local theory. The efforts to achieve this Bell test were awarded the 2022 Nobel prize in physics [8].

Clearly, incompatibility is an extraordinary feature. Yet, there is no single definition, approach, or even interpretation for it. In this dissertation alone there are eight approaches showcased, and surely others were left out. Historically, the main arguments for incompatibility rested on commutativity, based on the commutation between observables $[A, B] = AB - BA$, and uncertainty relations, which have been generalized from Heinserberg's proposal to any two observables [9], where, in this form, the uncertainty in the statistical dispersion of two observables is related to their commutativity. Because of this, many textbooks convolute both definitions [10–12].

However, the physical interpretation surrounding the uncertainty relation is very nuanced and subject to a lot of discussions [13, 14]. One of the possible interpretations relates the uncertainty to the measurement process, where it is understood that when a lower bound is present, it is impossible to perform the simultaneous measurements of the two observables in the same system without disturbing it. To formally define when observables can and cannot be simultaneously measurable, it is necessary to define unsharp observables, which carry intrinsic uncertainty in the form of “fuzziness” [15].

Such unsharp observables form a class of measurements with relaxed restrictions called Positive Operator-Valued Measurements (POVMs) [16]. Hence, the so-called *joint measureability* definition for incompatibility [17, 18] is formulated based on if a single device can produce the same outcomes simultaneously as two (or more) observables would. This definition has close ties to steering, a quantum feature that allows a distant part of the quantum system to non-locally influence (steer) the other part [19] – so close that, in fact, it is possible to define that if two observables can be used to demonstrate steering, they are necessarily not jointly measurable [20, 21].

Besides joint measurability, there has been another formulation for the interpretation of uncertainty arising in the measurement process, but now focused on the disturbance part of the interpretation: one cannot perform a measurement of an observable without disturbing the measurement of another incompatible observable on the same system. This description of incompatibility is called *nondisturbance* [22], and requires the concept of a quantum instrument.

Since an instrument implements the measurement of an observable onto an input state and gives an un-normalized state as the output, this allows for a definition based on if the measurement of a POVM, implemented through an instrument, can alter the probability distribution of another POVM. Therefore, one can identify the disturbance, or lack thereof, caused by the measurement of an observable. This definition is easily extendable for more than two observables, as long as each measurement implemented does not alter the probability distribution of the next one.

The definitions mentioned so far have always been stated in terms of incompatible observables, either generalized POVMs or not, implemented through instruments or not, and thus are accordingly categorized as Measurement Incompatibility and will be presented in more detail in Part I.

When one speaks of incompatibility, though, the measurements must be performed onto some state. If this state was, for example, heavy bodies submitted to thermal baths, measurements nearly cannot disturb them, even if they are deemed incompatible. It is reasonable to infer, then, that the state has an influence on the incompatibility status of the measurements conducted. This is what is argued in [23] when the authors propose a new approach to incompatibility: including the state in the definition. Not just as an input, but as an integral part of the definition. This is achieved through the context $\mathbb{C} = \{\rho, A, B\}$, which is the set of the measurements, $A = \sum_i^d a_i A_i = \sum_i^d a_i |a_i\rangle \langle a_i|$ and $B = \sum_j^d b_j B_j = \sum_j^d b_j |B_j\rangle \langle B_j|$, <u>and</u> the state ρ. Therefore, now it is the whole context that is incompatible or not.

This new approach was then dutifully followed by the proposal of a generalization based on relaxing the measurements to POVMs [24]. This attempt succeeded in providing a more generalized leakage detection protocol but resulted in the description now not necessarily claiming compatibility for classically behaving systems (like the heavy body) described through maximally mixed density operators. Both the original and the generalized measurement definitions are thus categorized as Context Incompatibility and will be presented in Part II of this dissertation.

Quantum theory in its “weirdness” has been criticized extensively. Many people were uncomfortable with the interpretations proposed and state of the art of the theory, like Einstein himself externalized in the aforementioned EPR paradox paper. Some proposed different interpretations that have an impact to this day, like the one in Ref. [25]. The experimental success of quantum theory as well as its many open questions (the measurement problem, quantum gravitation, etc.) have overall led to a search for basing the quantum theory on fewer postulates and more fundamental notions, modeling after the relativity theory. In 1994, for example, it was noticed by [26] in a paper focused on axiomatizing nonlocality that correlations respecting no-signaling theorems (i.e., no signal can travel faster than light, a desired feature as to not break the experimentally validated relativity theory) can violate Bell-type inequalities in even greater quantities than if they were also restricted by quantum mechanics. Such condition of no-signaling is also sufficient to produce stronger than quantum steering [27] and stronger than quantum key distribution protocols [28], related to cryptography tasks.

Therefore, other than ever so increasingly generalizing the measurements, we were compelled to search for a definition of context incompatibility that is entirely independent of the quantum mechanical framework. Based on probabilities and a simple idea of non-selective measurements alone, the proposed theory-independent definition would not only allow to brace for post-quantum theories that resolve quantum mechanics open questions but could usher the definition of incompatibility to a more fundamental one. Thus, we present this proposed definition in Part III, followed by the case study of the qubit (a two-level quantum system) in the Bloch representation and ending with the construction of quantifiers both for quantum contexts and for generalized probabilistic theories, but showcasing the evaluations mainly for the qubit.

1 MATHEMATICAL BACKGROUND

In this chapter, we will go through the mathematical background necessary for comprehending the details presented in this dissertation. For this purpose, we begin with the fundamentals of quantum mechanics, starting with the state vector and ending with the Bloch representation. Then we conclude with the concepts of classical probability theory that might be useful to bear in mind. This is intended as a short review, mainly to organize necessary notation and formulae for the future.

1.1 QUANTUM MECHANICS

This portion of the review is based mainly on the references [29–34].

1.1.1 State and Operators

In quantum mechanics, the state of a system is represented as a vector in the Hilbert space $\mathcal{H}$ through a *ket*, $|\psi\rangle$. In this work, we will be using only discrete d-dimensional Hilbert spaces, which have a one-to-one relation to the d-dimensional complex number field, i.e., $\mathcal{H} \simeq \mathbb{C}^d$, where $d > 1 \in \mathbb{N}$. In this representation, a system will often be best described by a linear combination of vectors, the famous superposition states, which can be written as

$$|\psi\rangle = \sum_i^d c_i |\psi_i\rangle, \tag{1.1}$$

where the $c_i \in \mathbb{C}$ must obey $\sum_i^d |c_i|^2 = 1$, stemming from the orthonormality relation of the basis[1], $\langle\psi_i|\psi_j\rangle = \delta_{ij}$.

To describe generally a system that is composed of more than one distinct physical system, for example, composed of a system $\mathcal{A}$ described by a vector in the basis $\{|\psi_i\rangle\}$ of $\mathcal{H}_\mathcal{A}$ and a system $\mathcal{B}$ described by a vector in the basis $\{|\varphi_j\rangle\}$ of $\mathcal{H}_\mathcal{B}$, one uses the tensor product:

$$|\phi\rangle = \sum_i^{d_\mathcal{A}} \sum_j^{d_\mathcal{B}} c_{ij} |\psi_i\rangle \otimes |\varphi_j\rangle \quad \in \quad \mathcal{H} = \mathcal{H}_\mathcal{A} \otimes \mathcal{H}_\mathcal{B} \tag{1.2}$$

where $d_{\mathcal{A}(\mathcal{B})}$ are the dimensions of the respective Hilbert spaces and c_{ij} are coefficients. The notations suppressing the tensor product $|\psi_i\rangle |\varphi_j\rangle$, $|\psi_i \varphi_j\rangle$ and variations like $|\psi_i\rangle_\mathcal{A} |\varphi_j\rangle_\mathcal{B}$ are also used. If c_{ij} can be decomposed as $c_{ij} = c_i c_j$, the state is called separable. If not, the state is entangled.

[1] The notation of the inner product in the Hilbert space is called *braket* and has the form $\langle\cdot|\cdot\rangle$, where $\langle\cdot| \in \mathcal{H}^*$, $\mathcal{H}^*$ being the dual space of $\mathcal{H}$, is the *bra*, essentially the transpose conjugate of $|\cdot\rangle$.

Other important elements in the quantum formalism are *linear operators* A that act on the Hilbert space transforming the vectors therein, such that

$$A\,|\psi\rangle = |\psi'\rangle\,, \quad A \in \mathcal{L}(\mathcal{H}), \tag{1.3}$$

where $\mathcal{L}(\mathcal{H})$ is the set of all bounded linear operators acting on $\mathcal{H}$. If this operator is Hermitian, that is, if its adjoint (or dual) operator $A^\dagger$ satisfies $A^\dagger = A$, it admits general spectral decomposition of the form

$$A = \sum_i^d a_i\,|a_i\rangle\,\langle a_i|\,, \tag{1.4}$$

where $a_i \in \mathbb{R}$ and $|a_i\rangle \in \mathcal{H}$ are the operator's eigenvalues and eigenvectors, respectively.

An important class of operators is the so-called projective operators, which, as the name suggests, operationalizes the orthogonal projection of a vector and have the general form

$$B_j = |b_j\rangle\,\langle b_j|\,, \tag{1.5}$$

so that when they act on a state $|\psi\rangle$ they project it to $|b_j\rangle$:

$$B_j\,|\psi\rangle = |b_j\rangle\,\underbrace{\langle b_j|\psi\rangle}_{\text{number}} = \langle b_j|\psi\rangle\,|b_j\rangle\,. \tag{1.6}$$

A set of projectors need to satisfy orthogonality, $\langle b_k|b_j\rangle = \delta_{jk}$, idempotency, $B_j^2 = B_j$, and completeness, $\sum_j^d B_j = \mathbb{1}$. They are, of course, also Hermitian.

Thereby, we can define a useful operation that acts on operators: the trace. The trace essentially takes only the diagonal terms of the matrix representation of an operator, and sums over all of them, that is,

$$\mathrm{Tr}[A] := \sum_k^d \langle a_k|\,A\,|a_k\rangle = \sum_{ki}^d a_i\,\langle a_k|a_i\rangle\,\langle a_i|a_k\rangle = \sum_{ki}^d a_i\delta_{ki} = \sum_i^d a_i, \tag{1.7}$$

where we conveniently chose the base that diagonalizes A to perform the trace, which is acceptable because the trace is base invariant. Besides this, another useful property is that the trace permits cyclic permutation ($\mathrm{Tr}[ABC] = \mathrm{Tr}[BCA] = \mathrm{Tr}[CAB]$).

Aside from the vector $|\psi\rangle$, there is a more generic representation for quantum systems in the form of *density operators*. These objects allow us to include our ignorance of the state itself, in the sense that we can encode a statistical description that accounts for the probability p_i of the system being represented by the specific $|\psi_i\rangle$ to the notation, which is thus given by

$$\rho = \sum_i^d p_i\,\underbrace{|\psi_i\rangle\,\langle\psi_i|}_{\rho_i} = \sum_i^d p_i\rho_i \in \mathcal{T}(\mathcal{H}), \tag{1.8}$$

where $p_i \in \mathbb{R}$ are classical weights and $\mathcal{T}(\mathcal{H}) \subseteq \mathcal{L}(\mathcal{H})$ is the set of trace-class operators acting on $\mathcal{H}$, which is thus named because $\mathrm{Tr}[\rho] = \sum_i^d p_i = 1$. Note that, since for projective operators $\mathrm{Tr}[B_j] = 1$, $B_j \in \mathcal{T}(\mathcal{H})$.

The properties that distinguish the density operator from any other operator are:

1. $\rho^\dagger = \rho$ (Hermiticity);
2. $\mathrm{Tr}[\rho] = 1$ (unitary trace);
3. $\langle\psi|\,\rho\,|\psi\rangle \geq 0 \quad \forall\,|\psi\rangle \neq 0$ (positivity).

If $\mathrm{Tr}[\rho^2] = 1$, the density operator is also idempotent and we call it *pure*, otherwise, we call it *mixed*. The reasoning for the names is that for pure operators $p_i = 1$, therefore there is no probability distribution of states, ρ being purely described by one $|\psi\rangle$. Whereas for mixed operators we have a convex sum of states to describe the system, a.k.a. a mixture of states.

Another important aspect of operators that is also valid for density operators is that one may be able to diagonalize them. The form of the diagonalized density operator is the spectral decomposition

$$\rho = \sum_k^d \lambda_k \,|\varphi_k\rangle \langle\varphi_k|\,, \tag{1.9}$$

where $\lambda_k \in \mathbb{R}$ is then the respective eigenvalues of the orthogonal basis $|\varphi_k\rangle$.

The description through density operators also allows for composite systems. Taking, for example, the composite system (1.2), the correspondent density operator is

$$\begin{aligned} \rho_{\mathcal{A},\mathcal{B}} =&\ |\phi\rangle\langle\phi| = \left(\sum_{ij} c_{ij}\,|\psi_i\rangle\,|\varphi_j\rangle\right)\left(\sum_{kl} c_{kl}^*\,\langle\psi_k|\,\langle\varphi_l|\right) \\ =&\ \sum_{ijkl} c_{ij}c_{kl}^*\,|\psi_i\rangle\langle\psi_k| \otimes |\varphi_j\rangle\langle\varphi_l| \quad \in \quad \mathcal{T}(\mathcal{H} = \mathcal{H}_{\mathcal{A}} \otimes \mathcal{H}_{\mathcal{B}}), \end{aligned} \tag{1.10}$$

which is a pure density operator, since $\rho_{\mathcal{A},\mathcal{B}}$ can be purely described by $|\phi\rangle$. On the other hand, a mixed composed density operator can be written as

$$\rho_{\mathcal{A},\mathcal{B}} = \sum_{ijkl} p_{ijkl}\,|\psi_i\rangle\langle\psi_k| \otimes |\varphi_j\rangle\langle\varphi_l|\,, \tag{1.11}$$

where p_{ijkl} cannot be decomposed as $c_{ij}c_{kl}^*$, for example. The trace operation on composite systems also have the properties of base invariance, cyclic permutation and unitary trace. It is defined as:

$$\mathrm{Tr}[\rho_{\mathcal{A},\mathcal{B}}] = \sum_{mn} \langle a_m| \otimes \langle b_n|\,\rho_{\mathcal{A},\mathcal{B}}\,|a_m\rangle \otimes |b_n\rangle\,, \tag{1.12}$$

where $\{|a_m\rangle\}(\{|b_n\rangle\})$ are any basis of $\mathcal{H}_{\mathcal{A}}(\mathcal{H}_{\mathcal{B}})$ for the trace.

There is an additional definition for composite systems, though: the partial trace. The partial trace is defined as taking the trace over only a part of the system. For example, the partial trace over part $\mathcal{A}$ of $\rho_{\mathcal{A},\mathcal{B}}$ is

$$\begin{aligned}
\mathrm{Tr}_{\mathcal{A}}[\rho_{\mathcal{A},\mathcal{B}}] &= \sum_m \langle\psi_m|\,\rho_{\mathcal{A},\mathcal{B}}\,|\psi_m\rangle = \sum_{mijkl} p_{ijkl}\,\langle\psi_m|\,(|\psi_i\rangle\,\langle\psi_k|\otimes|\varphi_j\rangle\,\langle\varphi_l|)\,|\psi_m\rangle \\
&= \sum_{mijkl} p_{ijkl}\,\underbrace{\langle\psi_m|\psi_i\rangle}_{\delta_{mi}}\,\underbrace{\langle\psi_k|\psi_m\rangle}_{\delta_{km}}\otimes|\varphi_j\rangle\,\langle\varphi_l| = \sum_{mjl} p_{mjml}\,|\varphi_j\rangle\,\langle\varphi_l| \\
&= \rho_{\mathcal{B}},
\end{aligned} \tag{1.13}$$

where the basis $\{|\psi_i\rangle\}$ for the trace was chosen for simplicity and the end result of a partial trace is shown to be a density operator only on part $\mathcal{B}$. Likewise,

$$\mathrm{Tr}_{\mathcal{B}}[\rho_{\mathcal{A},\mathcal{B}}] = \rho_{\mathcal{A}}. \tag{1.14}$$

With respect to the partial traces, the total trace can be written as

$$\mathrm{Tr}[\rho_{\mathcal{A}B}] = \mathrm{Tr}_{\mathcal{A}}[\mathrm{Tr}_{\mathcal{B}}[\rho_{\mathcal{A}B}]] = \mathrm{Tr}_{\mathcal{B}}[\mathrm{Tr}_{\mathcal{A}}[\rho_{\mathcal{A}B}]]. \tag{1.15}$$

1.1.2 Measurement

The measurement process in quantum mechanics is described through the action of Hermitian operators, then called *observables*, on states. The only possible results of a measurement $A = \sum_i^d a_i A_i$, where $A_i = |a_i\rangle\,\langle a_i|$ are projectors, upon the state $\rho = \sum_k p_k\,|\psi_k\rangle\,\langle\psi_k|$ are the eigenvalues of the said observable, and each eigenvalue occurs with probability given by

$$p_\rho(a_i) = \mathrm{Tr}[A_i\rho] = \langle a_i|\,\rho\,|a_i\rangle = \sum_k p_k|\,\langle a_i|\psi_k\rangle\,|^2. \tag{1.16}$$

The state immediately after a measurement is performed collapses to

$$\rho \to \rho_{a_i} = \frac{A_i\rho A_i}{\mathrm{Tr}[A_i\rho]}, \tag{1.17}$$

where the denominator $\mathrm{Tr}[A_i\rho]$ ensures normalization.

This is the standard presentation of quantum measurements, called Projective-Valued Measurement (PVM). It is possible, however, to define generalized measurements to require less restrictive operators than the projective ones. Namely, relaxing the requirement of orthogonality and Hermiticity, the set of operators $\{M_m\}$ needs only to have a completeness relation

$$\sum_m M_m^\dagger M_m = \mathbb{1}. \tag{1.18}$$

With these operators, it is possible to describe the probability of obtaining the outcome m when measuring in a system prepared in the state ρ:

$$p_\rho(m) = \mathrm{Tr}[M_m \rho M_m^\dagger], \tag{1.19}$$

which collapses the state to

$$\rho \to \rho_m = \frac{M_m \rho M_m^\dagger}{\mathrm{Tr}[M_m \rho M_m^\dagger]}. \tag{1.20}$$

Through the cyclic property of the trace, it is possible to write the probability as $\mathrm{Tr}[M_m^\dagger M_m \rho]$ and then define

$$E_m = M_m^\dagger M_m, \tag{1.21}$$

which are called the *effects*, clearly satisfying the completeness relation but also being Hermitian and positive, thus defining the more general Positive Operator-Valued Measurement (POVM) $\mathbb{E} = \{E_m\}_{m \in \Omega_E}$, where Ω_E is the collection of all possible measurement outcomes. Note that, if one requires that $M_m^\dagger = M_m$, $M_m^2 = M_m$, it all reduces to the projective measurements definitions.

In fact, according to Neumark's theorem [35], one realizes a POVM through a projective measurement onto an auxiliary system, also known as ancilla, entangled with the system. This means that instead of directly measuring the system $\mathcal{A}$, one first entangles it with an ancilla $\mathcal{B}$ and measures the ancilla with a PVM. From $\mathcal{A}$'s point of view, this is the same as performing a POVM.

Proof (finite case). Consider the system $\rho_\mathcal{A}$ and the ancilla $|0\rangle_\mathcal{B} \langle 0|_\mathcal{B}$. To entangle both systems, evolve the combined system with a joint unitary transformation $\mathrm{U}_{\mathcal{AB}}$ as follows

$$\mathrm{U}_{\mathcal{AB}} \left(\rho_\mathcal{A} \otimes |0\rangle_\mathcal{B} \langle 0|_\mathcal{B}\right) \mathrm{U}_{\mathcal{AB}}^\dagger = \mathrm{U}_{\mathcal{AB}}\, \rho_0\, \mathrm{U}_{\mathcal{AB}}^\dagger = \rho_t, \tag{1.22}$$

Then, perform a projective measurement $K = \sum_i k_i K_i = \sum_i k_i |k_i\rangle \langle k_i|$ onto the $\mathcal{B}$ part $P_i^\mathcal{B} = \mathbb{1}_\mathcal{A} \otimes K_i^\mathcal{B}$. The probability of obtaining outcome k_i is

$$\begin{aligned} p_{\rho_t}(k_i) &= \mathrm{Tr}[P_i^\mathcal{B} \rho_t] = \mathrm{Tr}\left[\mathbb{1}_\mathcal{A} \otimes K_i^\mathcal{B} \left(\mathrm{U}_{\mathcal{AB}}\, \rho_0\, \mathrm{U}_{\mathcal{AB}}^\dagger\right)\right] \\ &= \mathrm{Tr}\left[\mathbb{1}_\mathcal{A} \otimes |k_i\rangle_\mathcal{B} \langle k_i|_\mathcal{B}\, \mathrm{U}_{\mathcal{AB}} \left(\rho_\mathcal{A} \otimes |0\rangle_\mathcal{B} \langle 0|_\mathcal{B}\right) \mathrm{U}_{\mathcal{AB}}^\dagger\right] \\ &= \mathrm{Tr}_\mathcal{A}\left[\langle 0|_\mathcal{B}\, \mathrm{U}_{\mathcal{AB}}^\dagger |k_i\rangle_\mathcal{B} \langle k_i|_\mathcal{B}\, U_{\mathcal{AB}} |0\rangle_\mathcal{B}\, \rho_\mathcal{A}\right], \end{aligned} \tag{1.23}$$

where $\mathrm{Tr}_\mathcal{A}$ is the partial trace over $\mathcal{A}$ and the cyclic property of the trace was used. Define $M_i = \langle k_i|_\mathcal{B}\, U_{\mathcal{AB}} |0\rangle_\mathcal{B}$, which implies $M_i^\dagger = \langle 0|_\mathcal{B}\, \mathrm{U}_{\mathcal{AB}}^\dagger |k_i\rangle_\mathcal{B}$, and $E_i = M_i^\dagger M_i$, operators acting only in $\mathcal{B}$, to get

$$p_{\rho_t}(k_i) = \mathrm{Tr}_\mathcal{A}[M_i^\dagger M_i \rho_\mathcal{A}] = \mathrm{Tr}_\mathcal{A}[E_i \rho_\mathcal{A}]. \tag{1.24}$$

Since

$$\sum_i M_i^\dagger M_i = \sum_i \langle 0|_\mathcal{B} \mathrm{U}_{\mathcal{AB}}^\dagger \underbrace{|k_i\rangle_\mathcal{B} \langle k_i|_\mathcal{B}}_{\mathbb{1}} U_{\mathcal{AB}} |0\rangle_\mathcal{B}$$
$$= \sum_i \langle 0|_\mathcal{B} \underbrace{\mathrm{U}_{\mathcal{AB}}^\dagger U_{\mathcal{AB}}}_{\mathbb{1}} |0\rangle_\mathcal{B} = \sum_i \langle 0|0\rangle_\mathcal{B} = \mathbb{1}, \tag{1.25}$$

this means (1.23) corresponds to the probability (1.19) for a POVM $\mathbb{E} = \{E_i\}_{i\in\Omega_E}$ measurement onto the state $\rho_\mathcal{A}$: $p_{\rho_\mathcal{A}}(i)$.

Moreover, after the measurement is done, the state ρ_t collapses to

$$\frac{P_i^\mathcal{B}\rho_t P_i^\mathcal{B}}{p_{\rho_t}(k_i)} = \frac{\mathbb{1}_\mathcal{A} \otimes |k_i\rangle_\mathcal{B} \langle k_i|_\mathcal{B} \ \mathrm{U}_{\mathcal{AB}} \left(\rho_\mathcal{A} \otimes |0\rangle_\mathcal{B} \langle 0|_\mathcal{B}\right) \mathrm{U}_{\mathcal{AB}}^\dagger \ \mathbb{1}_\mathcal{A} \otimes |k_i\rangle_\mathcal{B} \langle k_i|_\mathcal{B}}{\mathrm{Tr}_\mathcal{A}[\rho_\mathcal{A} M_i^\dagger M_i]}$$
$$= \frac{\langle k_i|_\mathcal{B} U_{\mathcal{AB}} |0\rangle_\mathcal{B} \ \rho_\mathcal{A} \ \langle 0|_\mathcal{B} \mathrm{U}_{\mathcal{AB}}^\dagger |k_i\rangle_\mathcal{B}}{\mathrm{Tr}_\mathcal{A}[\rho_\mathcal{A} M_i^\dagger M_i]} \otimes |0\rangle_\mathcal{B} \langle 0|_\mathcal{B}$$
$$= \frac{M_i \rho_\mathcal{A} M_i^\dagger}{\mathrm{Tr}_\mathcal{A}[\rho_\mathcal{A} M_i^\dagger M_i]} \otimes |0\rangle_\mathcal{B} \langle 0|_\mathcal{B}, \tag{1.26}$$

which, for part $\mathcal{A}$, is exactly what is expected after a POVM measurement, as shown in (1.20), concluding the proof that a projective measurement on an entangled ancilla $\mathcal{B}$ corresponds to a POVM measurement on $\mathcal{A}$. ■

The proof given is a simplified version of the theorem's complete general proof, but it is enough to illustrate it for the purposes of this work.

1.1.3 Quantum Entropy

Entropy is one of those physical concepts that every physicist asked will answer a different definition, usually dependent on their field of study. As this work is inserted in the quantum field, we shall present here the quantum entropy definition, proposed by von Neumann (from the original article [36], translated book in [37]). It states that the entropy of a quantum state $\rho = \sum_k^d \lambda_k |\varphi_k\rangle \langle\varphi_k|$ is

$$S(\rho) := -\mathrm{Tr}[\rho \log \rho] = -\sum_k^d \lambda_k \log \lambda_k, \tag{1.27}$$

where the λ_k are the eigenvalues of ρ. Here, it is defined that $0 \log 0 \equiv 0$, based on the limit of the entropy for probabilities going to zero for $\lim_{\lambda_k\to 0} \lambda_k \log \lambda_k = 0$.

The quantum entropy (or von Neumann entropy) is interpreted as a measure of the observer's ignorance about the quantum state of the system since, when a system is exactly determined, i.e., it is a pure state, its entropy is $S(\rho) = 0$, aligning well with the interpretation that then there is no subjective ignorance regarding the state. On the other hand, if the system has an equal chance of being in each of the possible states, i.e., it is maximally mixed with $\lambda_k = 1/d$, its entropy reaches the maximum value $\log d$, aligning with the interpretation that the observer has maximum ignorance about which of the possible states such system is in.

Some very useful properties of this entropy are:

1. Non-negativity: $S(\rho) \geq 0$, equality holding iff ρ is a pure state;
2. Upper bound: $S(\rho) \leq \log d$, equality holding iff ρ is maximally mixed;
3. Invariance under unitary transformations: if $\mathrm{U}\mathrm{U}^\dagger = \mathrm{U}^\dagger\mathrm{U} = 1$, then $S(\mathrm{U}\rho\mathrm{U}^\dagger) = S(\rho)$;
4. Concavity: if $\rho = \sum_k^d \lambda_k \rho_k$, then $S(\sum_k^d \lambda_k \rho_k) \geq \sum_k^d \lambda_k S(\rho_k)$, equality holds if all the ρ_k are the same.

Another essential definition is the von Neumann relative entropy, which is related to how distinguishable two quantum states are, given by

$$\begin{aligned} S(\rho\|\sigma) :=\ & \mathrm{Tr}[\rho \log \rho] - \mathrm{Tr}[\rho \log \sigma] = -S(\rho) - \mathrm{Tr}[\rho \log \sigma] = -S(\rho) - \sum_j^d \langle \varphi_j | \rho \log \sigma | \varphi_j \rangle \\ =& -S(\rho) - \sum_j^d \langle \varphi_j | \rho | \varphi_j \rangle \log s_j = -S(\rho) - \sum_{jk}^d | \langle \sigma_j | \psi_k \rangle |^2 \lambda_k \log s_j, \end{aligned} \tag{1.28}$$

where $\rho = \sum_k^d \lambda_k |\psi_k\rangle \langle \psi_k|$ and $\sigma = \sum_j^d s_j |\varphi_j\rangle \langle \varphi_j|$ are the spectral decomposition of the states. The relative entropy is also non-negative, being zero iff $\rho = \sigma$.

1.1.4 Non-selective Measurement

An important scenario for this work is in the case where a projective measurement is performed but the outcome of the measurement is not recorded. This can be understood as performing the measurement but summing for all the possible results, collapsing the state to

$$\rho' = \sum_i^d A_i \rho A_i = \sum_i^d \langle a_i | \rho | a_i \rangle |a_i\rangle \langle a_i| = \sum_i^d \mathrm{Tr}[A_i \rho] A_i = \sum_i^d p_\rho(a_i) A_i, \tag{1.29}$$

which is a non-selective measurement. One can construct this by putting ρ through a non-selective quantum measurement map of the form

$$\Phi_A(\rho) = \sum_i^d A_i \rho A_i. \tag{1.30}$$

This map is completely positive trace-preserving (CPTP) and unitary, satisfying the necessary properties for one:

1. It can be written as $\sum_n^d K_n(\cdot)K_n^\dagger$, with K_n being Kraus operators [2] (implying positivity);

2. $\sum_n^d K_n^\dagger K_n = \mathbb{1}$ (implying trace preservation);

3. $\Phi_A(\mathbb{1}) = \mathbb{1}$ (unitary).

The map clears the state from any quantum coherence (the presence of off-diagonal terms) in its basis since it transforms the input state ρ into (1.29), which is in the diagonal basis of A. It characterizes a system where there is no information on the exact result of the measurement and as such, applying a map to the system can only increase its entropy:

$$S(\Phi_A(\rho)) \geq S(\rho). \tag{1.31}$$

Proof. Since $S(\rho\|\Phi_A(\rho)) = -S(\rho) - \mathrm{Tr}[\rho \log \Phi_A(\rho)] \geq 0$ and $S(\rho) \geq 0$, $S(\Phi_A(\rho)) \geq 0$, it is enough to prove that $-\mathrm{Tr}[\rho \log \Phi_A(\rho)] = S(\Phi_A(\rho))$ to satisfy the relation. To do this, apply the completeness relation $\sum_i A_i = \mathbb{1}$, followed by the idempotency $A_i = A_i^2$ and finally the cyclic property of the trace:

$$\begin{aligned} -\mathrm{Tr}[\rho \log \Phi_A(\rho)] &= -\mathrm{Tr}\left[\sum_i A_i \rho \log \Phi_A(\rho)\right] = -\mathrm{Tr}\left[\sum_i A_i \rho \underbrace{\log \Phi_A(\rho) A_i}_{A_i \log \Phi_A(\rho)}\right] \\ &= -\mathrm{Tr}\left[\sum_i A_i \rho A_i \log \Phi_A(\rho)\right] = -\mathrm{Tr}[\Phi_A(\rho) \log \Phi_A(\rho)] \\ &= S(\Phi_A(\rho)), \end{aligned} \tag{1.32}$$

thus completing the proof. ■

One can make successive non-selective measurements by applying sequential maps:

$$\begin{aligned} \Phi_{BA}(\rho) \equiv \Phi_B(\Phi_A(\rho)) &= \sum_j^d B_j \Phi_A(\rho) B_j \\ &= \sum_{ji}^d B_j A_i \rho A_i B_j \\ &= \sum_{ji}^d \underbrace{\mathrm{Tr}[A_i \rho]}_{p_\rho(a_i)} \underbrace{\mathrm{Tr}[A_i B_j]}_{p(b_j|a_i)} B_j. \end{aligned} \tag{1.33}$$

This type of non-selective measurement highlights how performing a measurement changes the system in quantum mechanics – even if no outcome is recorded or read.

[2] Kraus operators have a completeness relation (seen in item 2) but are neither idempotent nor orthogonal, just like the $\{M_i\}$ operators used to describe POVMs [38].

1.1.5 Reality

One interesting relation regarding the non-selective measurement map introduced above is the definition of *reality* proposed by Bilobran and Angelo in [39]. More specifically, the authors define what it takes for a physical quantity to be an element of reality of a state:

Definition 1.1. *The observable* $A = \sum_i a_i A_i = \sum_i a_i |a_i\rangle\langle a_i| \in \mathcal{L}(\mathcal{H})$ *is an element of reality for a state* $\rho \in \mathcal{T}(\mathcal{H})$ *iff*

$$\Phi_A(\rho) = \rho. \tag{1.34}$$

This definition is based on the idea that after a measurement if performed, regardless of if the outcome is recorded or not, there is an element of reality for the quantity measured in the system. From this idea, it is straightforward to see that for a state such as (1.34), A is already an element of reality since a second measurement of A would not change the state, only reveal this pre-existing reality: $\Phi_A(\rho) = \Phi_A(\Phi_A(\rho)) = \sum_{ij} A_i A_j \rho A_j A_i = \sum_i A_i \rho A_i = \Phi_A(\rho)$.

A quantifier for the violation of the reality of an observable in a system was proposed in the form of

$$\mathfrak{I}_A(\rho) := S(\Phi_A(\rho)) - S(\rho) = S(\rho \| \Phi_A(\rho)), \tag{1.35}$$

where the relation $\mathrm{Tr}[\rho f(\Phi_A(\rho))] = \mathrm{Tr}[\Phi_A(\rho) f(\Phi_A(\rho))]$, with $f(\cdot)$ being any function, was used [32]. Since the relative entropy is non-negative, being only equal to zero if the elements being compared are the same, this insures that the only case where no violation of reality is declared is for $\Phi_A(\rho) = \rho$, where we then say that ρ is a state of A-reality, or that ρ is A-real.

1.1.6 The Bloch Representation

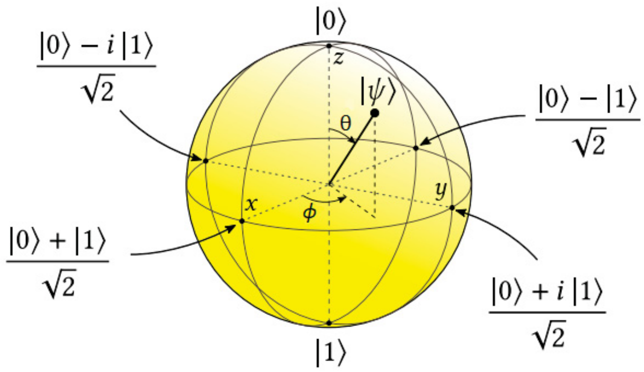
A practical representation for two-dimensional quantum systems is the Bloch sphere, which allows us to write operators as vectors on a three-dimensional sphere by parameterizing the state vectors coefficients by the angles $\theta \in [0, \pi]$ and $\phi \in [0, 2\pi]$.

A two-dimensional pure state, the qubit, can be generally written in the so-called computational basis as the vector

$$|\psi\rangle = a|0\rangle + b|1\rangle, \tag{1.36}$$

where $|a|^2 + |b|^2 = 1$ are the normalized coefficients. Therefore, we parameterize them as $a = \cos(\theta/2)$ and $b = e^{\imath\phi}\sin(\theta/2)$ and map the state in a sphere such as the one shown in Figure 1.

Figure 1 – Geometrical representation for the qubit, where it is possible to map state vectors as a vector therein. Illustration adapted from Ref. [34].



Writing (1.36) in the pure density operator format, one gets

$$\rho = |\psi\rangle\langle\psi| = \begin{pmatrix} |a|^2 & ab^* \\ a^*b & |b|^2 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1+\cos(\theta) & e^{-\imath\phi}\sin(\theta) \\ e^{\imath\phi}\sin(\theta) & 1-\cos(\theta) \end{pmatrix}, \tag{1.37}$$

where a^* and b^* are the complex conjugates of a and b. Since the Pauli matrices $\sigma = (\sigma_1, \sigma_2, \sigma_3)$, given by

$$\sigma_1 = \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \sigma_y = \begin{pmatrix} 0 & -\imath \\ \imath & 0 \end{pmatrix}, \quad \sigma_3 = \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{1.38}$$

alongside the identity matrix form a basis for a two-dimensional system, we can then write the density operator for a two-dimensional general mixed system as

$$\rho = \frac{1}{2}(\mathbb{1} + \mathbf{r}\cdot\sigma) = \frac{1}{2}\begin{pmatrix} 1+\mathbf{r}_z & \mathbf{r}_x - \imath\mathbf{r}_y \\ \mathbf{r}_x + \imath\mathbf{r}_y & 1-\mathbf{r}_z \end{pmatrix}, \tag{1.39}$$

where $\mathbf{r} = (r\sin[\theta]\cos[\phi], r\sin[\theta]\sin[\phi], r\cos[\theta])$ is a vector in the Bloch sphere such that $\|\mathbf{r}\| = \sqrt{\mathbf{r}\cdot\mathbf{r}} \in [0,1]$. It is clear from its form that if $\|\mathbf{r}\| = 0$, the state is maximally mixed corresponding to the central point of the Bloch sphere; if $\|\mathbf{r}\| = 1$, the state is pure (note that in this case, the state given in (1.39) would be equal to the one in (1.37)), lying on the shell of the sphere; and so any state for which $0 < \|\mathbf{r}\| < 1$ is simply a mixed state laying somewhere within the sphere.

On this same basis, it is possible to write projectors (1.5) as

$$B_j = \frac{1}{2}(\mathbb{1} + \mathbf{b}_j\cdot\sigma), \tag{1.40}$$

where the completeness relation fixes $\mathbf{b}_1 + \mathbf{b}_2 = \mathbf{0} \Rightarrow \mathbf{b}_1 = -\mathbf{b}_2$, so that we can write $\mathbf{b}_j = (-1)^j\hat{\mathbf{b}}$. A PVM $B = \sum_j b_j B_j$ with null trace (e.g., spin observables) also implies that the eigenvalues are -1 and 1, so that we can ultimately write the observable as $B = \hat{\mathbf{b}}\cdot\sigma$.

The projectors of the PVM rest on the shell of the Bloch sphere, thus forming a line crossing the sphere through the middle (a diameter) that represents the observable. Measuring in this setting means projecting the state vector on this line.

1.2 GENERIC PROBABILISTIC THEORIES

As the title of this dissertation suggests, we are interested in other not-necessarily quantum mechanic elements to achieve a theory-independent definition. Here we highlight the definition of probabilities and a classical version of the relative entropy.

1.2.1 Probabilities

Some notion of classical probability theory is required for this work, albeit nothing very rigorous. This Subsection is based mainly on the references [29, 32, 40, 41].

We say that a random variable X that can assume i values when measured in a system $\mathscr{E}$ has a sample space $\mathcal{S}$, consisting of all the possible x_i results. The possible values as well as combinations of them are called events, for example, if $X = \{1, 2, 3, ..., x_i\}$, each x_i is an event, but also, every even result can be an event or any grouping desired. The probability of a certain x_i outcome happening is, then, $p_{\mathscr{E}}(x_i)$, satisfying the following conditions

1. $0 \leq p_{\mathscr{E}}(x_i) \leq 1$;
2. $p_{\mathscr{E}}(\mathcal{S}) = 1$, equivalently written in the discrete case as $\sum_i^n p_{\mathscr{E}}(x_i) = 1$ or, sometimes, $p_{\mathscr{E}}(\{x_i\})$;
3. If the events E_i are mutually exclusive, the probability of at least one of them occurring is $p_{\mathscr{E}}(\bigcup_i E_i) = \sum_i p_{\mathscr{E}}(E_i)$.

These are the axioms of the modern probability theory, but there are some other important propositions:

1. For any two events A and B, $p_{\mathscr{E}}(A \cup B) = p_{\mathscr{E}}(A) + p_{\mathscr{E}}(B) - p_{\mathscr{E}}(A \cap B)$. The probability $p_{\mathscr{E}}(A \cap B)$ is commonly written as $p_{\mathscr{E}}(A, B)$, meaning the joint probability of the events happening together;
2. For three events A, B and C, $p_{\mathscr{E}}(A \cup B \cup C) = p_{\mathscr{E}}(A) + p_{\mathscr{E}}(B) + p_{\mathscr{E}}(C) - p_{\mathscr{E}}(A, B) - p_{\mathscr{E}}(A, C) - p_{\mathscr{E}}(B, C) + p_{\mathscr{E}}(A, B, C)$, and so on for more events;
3. If the events are independent, $p_{\mathscr{E}}(A, B) = p_{\mathscr{E}}(A)p(B)$;

4. The conditional probability of an event A occurring, given that the event B has occurred, is $p_{\mathcal{E}}(A|B) := \frac{p_{\mathcal{E}}(A,B)}{p_{\mathcal{E}}(B)}$.

These definitions and the common basic logic of probabilities are enough for the purpose of this dissertation, for more in-depth information we refer to [40, 41].

1.2.2 Divergence

Sometimes, it is important to describe how distinguishable are two probability distributions, $p_{\mathcal{E}}(\{x_i\})$ and $q_{\mathcal{E}}(\{x_i\})$, for the same variable. This can be done by means of the Kullback-Liebler divergence [42], which is the classical version of the relative entropy (1.28) and is given by

$$\begin{aligned} D(p_{\mathcal{E}}(\{x_i\})\|q_{\mathcal{E}}(\{x_i\})) &= \sum_i^d p_{\mathcal{E}}(x_i)\log\frac{p_{\mathcal{E}}(x_i)}{q_{\mathcal{E}}(x_i)} \\ &= \sum_i^d p_{\mathcal{E}}(x_i)\log p_{\mathcal{E}}(x_i) - \sum_i^d p_{\mathcal{E}}(x_i)\log q_{\mathcal{E}}(x_i) \\ &= -H(p_{\mathcal{E}}\{x_i\}) - \sum_i^d p_{\mathcal{E}}(x_i)\log q_{\mathcal{E}}(x_i) \geq 0, \end{aligned} \tag{1.41}$$

where $H(p\{x_i\}) = -\sum_i^d p(x_i)\log p(x_i) \geq 0$ is the Shannon entropy [43], related to the ignorance about the outcome of a variable, similarly to how the von Neumann entropy is also related to ignorance. In fact, Shannon entropy is a generalization of the von Neumann entropy. The divergence is equal to zero iff the two probability distributions are the same, i.e., $p_{\mathcal{E}}(\{x_i\}) = q_{\mathcal{E}}(\{x_i\})$.

Part I

Measurement Incompatibility

2 INCOMPATIBILITY DEFINITIONS

Many courses in quantum mechanics avoid discussing in depth the concept of incompatibility, with some textbooks only tip-toeing around the subject without even using the word “incompatible” [29]. Most of the time, the discussion is taken a step forward straight to the consequences of the incompatible nature of quantum mechanics. There is a good reason for that: the definition of incompatibility is not at all unanimous, the lack of a classical counterpart makes it so that even the interpretation surrounding it is not on common ground, and the mathematical tools for each definition vary from quite simple to very tortuous. Owing to that, this chapter is reserved for this pillar of quantum weirdness and we will approach the subject from a somewhat chronological order but reserving the right to discuss the simpler forms earlier.

2.1 COMMUTATION

The commutator is a relation between two operators acting on the same Hilbert space $\mathcal{H}$, A and $B \in \mathcal{L}(\mathcal{H})$, defined as

$$[A, B] \equiv AB - BA. \tag{2.1}$$

We say that A and B commute if $[A, B] = 0$ and, if two observables commute, we say they are *compatible*. Thus, if $[A, B] \neq 0$, A and B are *incompatible*.

This is a straightforward, raw, mathematical definition for the incompatibility of observables. However, the connection with our physical intuition is subtler, as the relation (2.1) does not form an observable (it's not Hermitian) and thus cannot be measured directly. For this discussion, consider the following theorem [29, 44]:

Theorem 2.1 (Simultaneous Diagonalization)**.** *Suppose A and B are two observables. Then $[A, B] = 0$ if there exists an orthonormal basis such that both A and B are diagonal with respect to that basis. We say that A and B are simultaneously diagonalizable in this case.*

Proof. Consider the compatible pair A and B. Now, take the orthonormal basis for which A is diagonal given by $|a_i\rangle$. We have $\langle a_j| [A, B] |a_i\rangle = 0$, but

$$\langle a_j| [A, B] |a_i\rangle = \langle a_j| AB |a_i\rangle - \langle a_j| BA |a_i\rangle = (a_j - a_i) \langle a_j| B |a_i\rangle = 0,$$

which implies that, for $a_i \neq a_j$, $\langle a_j| B |a_i\rangle = 0$. Therefore B is diagonal in the same orthonormal basis that diagonalizes A. Analogously, the same could be done taking first the orthonormal basis that diagonalizes B.

Now, consider that A and B are simultaneously diagonalizable in the basis $|a_i, b_i\rangle$ so that $A|a_i, b_i\rangle = a_i|a_i, b_i\rangle$ and $B|a_i, b_i\rangle = b_i|a_i, b_i\rangle$. This means that $AB|a_i, b_i\rangle = Ab_i|a_i, b_i\rangle = a_ib_i|a_i, b_i\rangle$ but also $BA|a_i, b_i\rangle = Ba_i|a_i, b_i\rangle = a_ib_i|a_i, b_i\rangle$, and therefore $AB = BA$. ■

In addition to the simultaneous diagonalization theorem, consider the collapse postulate of quantum mechanics [30]:

Postulate (Collapse)**.** *If the measurement of the observable $A = \sum_i a_i A_i$ on the system in a state $|\psi\rangle$ gives the result a_i, the state of the system immediately after the measurement is the normalized projection $|\psi\rangle \to |\psi'\rangle = \frac{A_i|\psi\rangle}{\sqrt{\langle\psi|A_i|\psi\rangle}}$ of $|\psi\rangle$ onto the eigensubspace associated with a_i.*

Hence, if two observables commute, the shared diagonal basis makes it so that applying either observable will collapse the state to a basis that allows for the attainment of the same eigenvalues for both observables. This indicates that the measurement could be performed regardless of order and simultaneously in the same system without any formal restriction.

The aforementioned non commutability also has an impact on the acquisition of information about the system. To enlighten this idea, take now, as an example, the qubit state $|\psi\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$. If one makes a spin measurement in the z direction, given by $S_z = \frac{\hbar}{2}(|0\rangle\langle 0| - |1\rangle\langle 1|)$, the state after obtaining the outcome $+\frac{\hbar}{2}$ would simply become the eigenstate of S_z: $|\psi'\rangle = |0\rangle$. If then one measures the non-commuting (incompatible) spin in the x direction, given by $S_x = \frac{\hbar}{2}(|+\rangle_x\langle +|_x - |-\rangle_x\langle -|_x)$, where $|\pm\rangle_x = \frac{1}{\sqrt{2}}(|0\rangle \pm |1\rangle)$, and obtained $-\hbar/2$ as a result, the state would become $|\psi''\rangle = |-\rangle_x = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$. This is rather odd because the superposition in the z direction has been restored, indicating that if one tried to measure again in this direction, obtaining $-\frac{\hbar}{2}$ as a result would be possible (it is, in fact, a possibility of 50%), even though one had previously obtained $+\frac{\hbar}{2}$ for the same system! We say that measuring the x-spin, in a way, destroys the information regarding the z-spin, which is then an expression of measurement incompatibility.

This interpretation of the incompatibility of measurements as information destruction is a very common one, as it can be derived from these two very fundamental elements of quantum mechanics (the Theorem 2.1 and the collapse postulate).

2.2 UNCERTAINTY RELATIONS

Historically, the Uncertainty Principle is derived considering the wave function of a free particle at $t = 0$, $\psi(x, 0) = \frac{1}{\sqrt{2\pi\hbar}}\int \bar{\psi}(p)e^{ipx/\hbar}dp$. With the form of the wave packet and the Fourier Transform relation with $\bar{\psi}(p)$, it is possible to achieve the relation between the widths of the wave packet, Δx and Δp:

$$\Delta x \Delta p \gtrsim \hbar. \tag{2.2}$$

This relation, commonly known as Heisenberg's Uncertainty Principle [1], tells us that there is a limit on the spreading of the wave packet with regard to the position and momentum. Since Δx and Δp are related to the width of the peak around x_0 and p_0, the initial positions and momentum, respectively, as we try to determine the point x_0 to a very narrow uncertainty, Δp needs to then increase in width. Or, as is usually put, one cannot determine both position and momentum of a particle to an arbitrary degree of accuracy. It is implied, then, that such a limit on the uncertainties is a physical restriction (or, at least, pertaining to the physical theories for which these mathematical conditions applies[1]) that cannot be overcome.

The uncertainty equation (2.2) can be generalized using the *braket* formalism of quantum mechanics [9] for any two observables acting on the same Hilbert space $\mathcal{H}$. For this, we define the operator

$$\Delta A \equiv A - \langle A \rangle, \tag{2.3}$$

where $\langle A \rangle = \langle \psi | A | \psi \rangle = \mathrm{Tr}[A\rho]$ is the expectation value of the observable A. We can then write the dispersion of A, also known as mean square deviation or variance, given by

$$\langle (\Delta A)^2 \rangle = \langle A^2 - 2A\langle A \rangle + \langle A \rangle^2 \rangle = \langle A^2 \rangle - \langle A \rangle^2. \tag{2.4}$$

Using these definitions alongside the Cauchy-Schwarz inequality [45, 46] with another observable B, we obtain

$$\langle (\Delta A)^2 \rangle \langle (\Delta B)^2 \rangle \geq |\langle \Delta A \Delta B \rangle|^2. \tag{2.5}$$

Noting that

$$\Delta A \Delta B = \frac{1}{2}[\Delta A, \Delta B] + \frac{1}{2}\{\Delta A, \Delta B\}, \tag{2.6}$$

where $[\Delta A, \Delta B] = [A, B]$ is the commutator (2.1), an anti-Hermitian operator which always has purely imaginary expectation value, and $\{\Delta A, \Delta B\} = \Delta A \Delta B + \Delta B \Delta A$ is the anti-commutator, which is Hermitian and has purely real expectation value. We have, then,

$$|\langle \Delta A \Delta B \rangle|^2 = \frac{1}{4}|\langle [A, B] \rangle|^2 + \frac{1}{4}|\langle \{\Delta A, \Delta B\} \rangle|^2. \tag{2.7}$$

[1] It is possible to derive relations between conjugate variables of this type for the waves in the Electromagnetic theory. In quantum mechanics, however, these wave functions carry information about particle behaviors, which makes such uncertainty relations "weirder", in the sense that classically we wouldn't expect material things to have limits on the certainty of position and momentum [30].

This allows us to write (2.5) as

$$\langle(\Delta A)^2\rangle\langle(\Delta B)^2\rangle \geq \frac{1}{4}|\langle[A,B]\rangle|^2 + \frac{1}{4}|\langle\{\Delta A, \Delta B\}\rangle|^2. \tag{2.8}$$

Often, though, authors like [44, 47, 48] will omit the term for the anti-commutator, leaving the relation as

$$\langle(\Delta A)^2\rangle\langle(\Delta B)^2\rangle \geq \frac{1}{4}|\langle[A,B]\rangle|^2. \tag{2.9}$$

The reasoning for this omission (when provided) is that it can only make the lower bound stronger, as in order to satisfy (2.8), surely (2.9) would have to be satisfied.

The uncertainty relations are interesting as they provide a lower bound for something that one can actually measure in the laboratory: the dispersion of the observables, which is simply the statistical dispersion associated with the several measurements of a said observable [29], as opposed to the commutator, which is not an observable at all. Both definitions of incompatibility are related, though, as it is obvious from the right-hand side of Equation (2.9) that if one has commuting observables, the lower bound is zero and then quantum states exist implying arbitrarily low dispersion (high determinacy) for such observables, having no destruction of information.

2.3 JOINT MEASUREABILITY

The physical interpretation of the uncertainty principle is highly debated [13, 14], usually reduced to two possibilities:

1. Uncertainty is related to the preparation of the system, which is the idea that it is impossible to prepare a state that has sharply localized both position probability distribution and momentum probability distribution;

2. Uncertainty is related to the measurement process, the idea that one cannot implement simultaneous measurements without disturbing the probability distributions, regardless of the state input.

The first interpretation is straightforward and to implement such a verification of the uncertainty on the preparation, one needs only to perform several rounds of measurement of each observable on copies of the same state to obtain the variance in the results. The second interpretation, though, requires the measurement of the observables in the same state and can be verified through the possibility of obtaining both statistical distributions with the same apparatus or through comparing the distributions of the results depending on the order of measurement. In the double-slit experiment, for example, this could mean comparing the distribution of position found normally with the distribution found if one tries to detect the passage (a detection of momentum) through one of the slits.

To explain the impossibility of simultaneous measurement on the second type of interpretation the idea that the observables themselves carried inherent uncertainty arose [15]. To represent this uncertainty within observables in the quantum formalism it is used the concept of unsharp observables, which carry this uncertainty in the form of a "fuzziness", having eigenvalues in a range of the interval $[0, 1]$ instead of the dichotomous 0 or 1 of projective observables [33].

For the definition of unsharp observable, consider first the PVM observable A where $A = \sum_{i \in \Omega_A} a_i A_i$, with Ω_A being the set of all measurement outcomes of A. We will, from now on, also refer to this type of projective measurement as a *sharp measurement*.

To define an unsharp observable we simply add a noise parameter (μ, for example) to the sharp observables in the form:

$$A_i^{\mu} \equiv (1-\mu)\frac{\mathbb{1}}{d} + \mu A_i, \quad \mu \in [0, 1], \tag{2.10}$$

where d is the dimension of the Hilbert space. Such noisy operator can be counted, actually, as the i-th effect of the generalized observable, the POVM $\mathbb{A} \equiv \{A_i^{\mu}\}_{i \in \Omega_A}$, introduced in Subsection 1.1.2.

With the notion of unsharp observable and POVMs, we can define joint measurability [33, 49]:

Definition 2.2. *Two POVMs $\mathbb{A} = \{A_i^{\mu}\}_{i \in \Omega_A}$ and $\mathbb{B} = \{B_j^{\nu}\}_{j \in \Omega_B}$ can be jointly measured if there exists a third POVM $\mathbb{G} = \{G_{i,j}^{\mu\nu}\}$ such that*

$$A_i^{\mu} = \sum_j G_{i,j}^{\mu\nu}, \quad \forall \mu, \quad B_j^{\nu} = \sum_i G_{i,j}^{\mu\nu}, \quad \forall \nu. \tag{2.11}$$

The POVM $\mathbb{G}$ is called a joint observable or parent POVM of $\mathbb{A}$ and $\mathbb{B}$.

Incompatible observables here, therefore, would be the set of POVMs that cannot be jointly measured. The interpretation of this definition resides in the idea that if observables are incompatible it is impossible to construct a single device that could implement both measurements simultaneously. In this structure, sharp observables are simply a special case for $\mu = 1$, and so the projective measurements are not lost.

It is important to notice that this construction allows for observables usually deemed incompatible to be jointly measured given enough noise is added. To illustrate this, consider the unsharp spin-z and spin-x observables:

$$S_{i|x}^{\mu} = \frac{1}{2}(\mathbb{1} + (-1)^i \mu \sigma_x), \tag{2.12}$$

$$S_{j|z}^{\mu} = \frac{1}{2}(\mathbb{1} + (-1)^j \mu \sigma_z), \tag{2.13}$$

where $0 < \mu \leq 1$ is the noise parameter and $i, j \in 0, 1$. A possible joint observable of them is

$$G^{\mu}_{i,j} = \frac{1}{4}(\mathbb{1} + (-1)^i \mu\sigma_x + (-1)^j \mu\sigma_z), \tag{2.14}$$

Clearly,

$$S^{\mu}_{i|x} = \sum_j G^{\mu}_{i,j}, \tag{2.15}$$

$$S^{\mu}_{j|z} = \sum_i G^{\mu}_{i,j}. \tag{2.16}$$

To completely fulfill definition (2.11), $G^{\mu}_{i,j}$ needs to be a POVM, thus, positive, which only happens for $\mu \leq 1/\sqrt{2}$. Hence, for noise parameters lower than (or equal to) $1/\sqrt{2}$, the noisy spin observables are jointly measureable and therefore compatible! This is also a good example of how two observables that do not commute can be approximately compatible (in the sense that their unsharp versions are approximations) according to this definition.

2.4 NONDISTURBANCE

In order to focus on the interpretation of measurements disturbing quantum systems, it is possible to formally define nondisturbance. This is also done for POVMs and includes the idea of sequential measurement, where one compares statistical distributions for an observable with and without the previous measurement of some other observable to then conclude on their ability to disturb the system [14].

Besides POVMs, in order to define nondisturbance it is necessary to introduce the concept of *instruments* [22]. An instrument which implements a POVM $\mathbb{A} = \{A_x\}_{x \in \Omega_A}$ is a collection of completely positive linear maps $\mathcal{I}_\Lambda(\rho) \equiv \{\mathcal{I}_x(\rho)\} \in \mathcal{T}(\mathcal{H})$, where $\mathcal{I}_x(\rho)$ is the un-normalized state after obtaining x as the measurement outcome when the instrument $\mathcal{I}_\Lambda$ implements $\mathbb{A}$ in the system state ρ. We say that an instrument $\mathcal{I}_\Lambda$ implements $\mathbb{A}$ when the condition

$$\mathcal{I}^{\dagger}_x(\mathbb{1}) = A_x, \quad \forall x \tag{2.17}$$

holds. The probability of the outcome x occurring is $\mathrm{Tr}[\mathcal{I}_x(\rho)] = \mathrm{Tr}[\rho A_x]$. If one ignores the measurement outcome, though, the instrument $\mathcal{I}_\Lambda$ transforms ρ into

$$\mathcal{I}_\Lambda(\rho) = \sum_x \mathcal{I}_x(\rho). \tag{2.18}$$

Notice that, since $\sum_x A_x = \mathbb{1}$, we have $\sum_x \mathrm{Tr}[\mathcal{I}_x(\rho)] = \sum_x \mathrm{Tr}[\rho A_x] = 1$, and so $\mathrm{Tr}[\mathcal{I}_\Lambda(\rho)] = 1$. Therefore, $\mathcal{I}_\Lambda(\rho)$ corresponds to a CPTP map, introduced in Subsection 1.1.4.

Clear from its formulation, many different instruments correspond to the same observable. An example of an instrument implementing A is the Lüders instrument [50], given by

$$\mathcal{I}_x^L(\rho) = A_x^{1/2}\rho A_x^{1/2}. \tag{2.19}$$

Now we are able to introduce the definition of nondisturbance.

Definition 2.3. *Given two POVMs $\mathbb{A} = \{A_x\}_{x\in\Omega_A}$ and $\mathbb{B} = \{B_y\}_{y\in\Omega_B}$, we say $\mathbb{A}$ can be measured without disturbing $\mathbb{B}$ if there exists an instrument that implements $\mathbb{A}$ and for which*

$$\mathrm{Tr}[\mathcal{I}_\Lambda(\rho)B_y] = \mathrm{Tr}[\rho B_y] \quad \forall \rho, y. \tag{2.20}$$

Equation (2.20) indicates that the measurement statistics of $\mathbb{B}$ are unaffected by the measurement of $\mathbb{A}$, since $\mathcal{I}_\Lambda(\rho)$ implements those measurements of $\mathbb{A}$ upon ρ. As an example, consider the commutative pair of observables $[A_x, B_y] = 0$ and the implementation of $\mathbb{A}$ via the Lüders Instrument, yielding

$$\mathrm{Tr}[\mathcal{I}_\Lambda^L B_y] = \mathrm{Tr}\left[\sum_x A_x^{1/2}\rho A_x^{1/2} B_y\right] = \sum_x \mathrm{Tr}[A_x\rho B_y] = \mathrm{Tr}[\rho B_y], \tag{2.21}$$

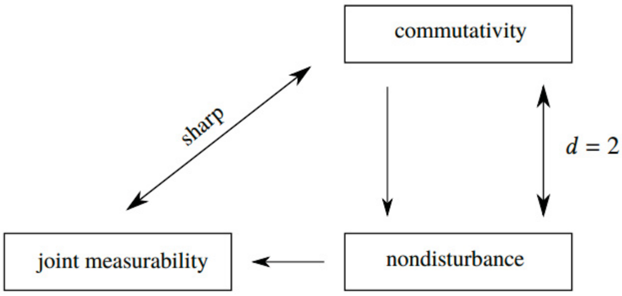
where the cyclic property of the trace and completeness property of A_x were used, showing that when the observables commute, one cannot disturb the other's statistics. In fact, this is true for every pair of observables, meaning that whenever two observables commute there is no instrument that could lead to a violation of (2.20). The converse is only true for two-dimensional systems.

The relationship showed between commutativity and nondisturbance can be explored for other definitions of incompatibility. For example, if an instrument implements $\mathbb{A}$ and does not disturb $\mathbb{B}$, then $\mathbb{A}$ and $\mathbb{B}$ are jointly measurable, as one could write a mother-POVM in the form of $G_{x,y} = \mathcal{I}_x^\dagger(B_y)$. Therefore, nondisturbance implies joint measurability, but the converse is not true.

From these two relationships follows a third one: if commutativity is respected, so is joint measurability. And lastly, if the measurements are sharp, all three definitions agree. All of these relationships are illustrated in Figure 2.

Bringing the discussion on nondisturbance to an end, it is straightforward to see the connection between nondisturbance and loss of information, as a possible interpretation of what constitutes the disturbance is precisely whether or not there is loss of information affecting the probability distribution of an observable if first it was implemented another observable on the state.

Figure 2 – Figure displaying the relationship between nondisturbance, joint measurability, and commutativity. The one-sided arrows indicate one-way relationships and the two-sided arrows indicate two-way relationships and their conditions. Diagram taken from Ref. [33].



2.5 TAKEAWAY MESSAGE

In this chapter we hope to have made the following points:

- Commutativity is well established as an incompatibility definition *for projective observables*. Albeit restrictive, observables that violate commutativity have been extensively used experimentally, like in the loophole-free Bell test [7];
- All of the definitions presented in this Section have some relationship with commutativity, being it explicit, like the uncertainty relations, or implicit by requiring restrictions to reduce to commutativity (see Figure 2);
- None of the incompatibility proposals put the state itself in high regard. The uncertainty relation may be dependent on the state, e.g., for the spin operators $[S_i, S_j] = \epsilon_{ijk} \imath\hbar S_k$, where ϵ_{ijk} is the Levi-Civita symbol[2], which implies that the uncertainty relation would depend on $\langle S_k \rangle = \mathrm{Tr}[S_k \rho]$, but that is not always the case, as for example, for $[X, P] = \imath\hbar$. The nondisturbance definition explicitly features the state, but the $\forall \rho$ requirement essentially nullifies its effect on the overall definition.

2 The Levi-Civita symbol is defined as $\epsilon_{ijk} = \begin{cases} +1 & \text{if } (ijk) \text{ is an even permutation,} \\ -1 & \text{if } (ijk) \text{ is an odd permutation,} \\ 0 & \text{otherwise.} \end{cases}$

Part II

Context Incompatibility

3 THE MSA CONTEXT COMPATIBILITY

After presenting many incompatibility definitions, the reader may have noticed a trend: the focus on the measurement, whether it is projective measurements or POVMs. This focus is understandable considering that early works such as Heisenberg's [1] noticed the uncertainty relation between position and momentum – no sight of the state itself on the final equation. However, works such as [23] have argued that considering the state may be more intuitive to the construction of an incompatibility notion since it's well-known that such an effect vanishes on macroscopic systems, e.g., heavy bodies. It seems reasonable, then, to require that for the same set of measurements a change in the system being measured from a quantum regime to a classical one suffices in making the measurements compatible. Ergo, a definition that included the state itself and not only the observables was proposed: the Context Incompatibility, which this Part is reserved for presenting.

3.1 CONTEXT, INFORMATION, AND PROTOCOL

A context is defined as the set $\mathbb{C} = \{\rho, A, B\} \subset \mathcal{L}(\mathcal{H})$ formed by a state $\rho \in \mathcal{T}(\mathcal{H})$ and observables $A = \sum_i^d a_i A_i = \sum_i^d a_i |a_i\rangle\langle a_i|$, $B = \sum_j^d b_j B_j = \sum_j^d b_j |b_j\rangle\langle b_j| \in \mathcal{L}(\mathcal{H})$, all acting on the same Hilbert space. Note that this is initially defined only for projective observables, instead of the generalized POVMs. This is a choice made by the authors considering Neumark's Theorem (original [35], found in English in [51, 52]), which reads that any POVM can be realized by introducing an auxiliary system (also known as ancilla), performing a unitary transformation on the combined system and then making a projective measurement on the ancilla.

Consider now the protocol depicted in Figure 3, where Alice prepares a state ρ with information content given by

$$I(\rho) = \log d - S(\rho), \tag{3.1}$$

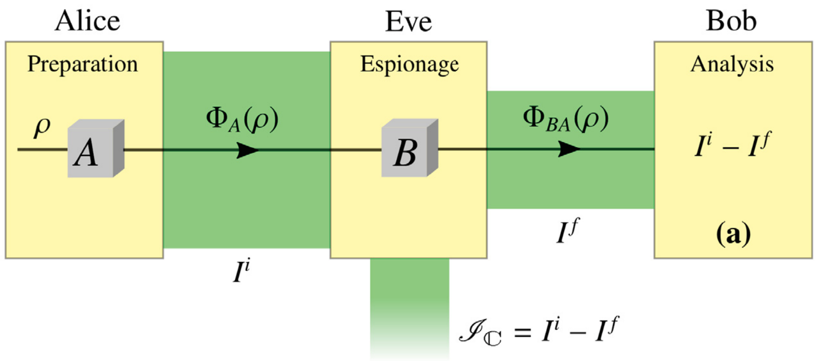
where $S(\rho)$ is the von Neumann entropy of ρ and $\log d$ is the maximum entropy possible for a d-dimensional state. After measuring A without registering the result, Alice transforms the state into

$$\Phi_A(\rho) = \sum_{i=1}^{d} A_i \rho A_i = \sum_{i=1}^{d} p_\rho(a_i) A_i, \tag{3.2}$$

where $p_\rho(a_i) = \mathrm{Tr}[\rho A_i]$ and $\Phi_A(\rho)$ is the non-selective measurement map (1.30). Consequently, Alice reduces the information content to

$$I^i \equiv I(\Phi_A(\rho)) = \log d - S(\Phi_A(\rho)) = \log d - H(\{p_\rho(a_i)\}), \tag{3.3}$$

Figure 3 – Depiction of the leakage detection protocol for a quantum communication channel, showcasing on the left the non-selective measurement made by Alice $\Phi_A(\rho)$ with information content I^i illustrated by the biggest green stripe. This is followed by the non-selective measurement on the middle that represents the interference of a spy on the system, Eve, collecting the information amount $\mathscr{I}_C$ (smallest green stripe) and sending out the state $\Phi_{BA}(\rho)$ with respective information content I^f, which Bob will receive on the right. Adapted figure from Ref. [23].



where $H(\{p_\rho(a_i)\}) = -\sum_i^d p_\rho(a_i)\log p_\rho(a_i)$ is the Shannon entropy of the probability distribution $\{p_\rho(a_i)\}$. Then, Alice sends this state to Bob, who expects to receive I^i information resource, as prearranged with Alice. That is, of course, if there is no information leakage.

Suppose now that there may be an eavesdropper in the system, Eve, attempting to steal information from the communication channel through the non-selective measurement of B. She might do it through a unitary transformation $\mathrm{U} \in \mathcal{L}(\mathcal{H} \otimes \mathcal{H}_\mathcal{E})$ that entangles her apparatus, $\mathcal{E}$, with the system that left Alice's laboratory, $\Phi_A(\rho)$. Ultimately, the difference between the information content arriving at (I^i) and leaving (I^f) Eve's lab, therein defined as the consumed information $\mathscr{I}_C$, is

$$\mathscr{I}_C := I^i - I^f = I(\Phi_A(\rho)) - I(\Phi_{BA}(\rho)) = S(\Phi_A(\rho)\|\Phi_{BA}(\rho)), \tag{3.4}$$

where the relation $\mathrm{Tr}[\rho f(\Phi_A(\rho))] = \mathrm{Tr}[\Phi_A(\rho) f(\Phi_A(\rho))]$, with $f(\cdot)$ being any function, was used to validate the relation $S(\Phi_A(\rho)) - S(\rho) = S(\rho\|\Phi_A(\rho))$ [32]. Hence, if Bob analyzed $\mathscr{I}_C$, he would discover the status of leakage on the channel. The context incompatibility definition based on this protocol is then as follows:

Definition 3.1. *Context incompatibility is the resource encoded in a context* $\mathbb{C} = \{\rho, A, B\}$ *that allows one to test the safety of a communication channel against information leakage. Quantified via* $\mathscr{I}_C = S(\Phi_A(\rho)\|\Phi_{BA}(\rho))$*, it is operationally related to the amount of information subtracted from the system upon an external measurement.*

If Bob finds $\mathscr{I}_C = 0$, the conclusion would be that there was no leakage on the channel. From the property of the relative entropy, we have that $\mathscr{I}_C = 0$ iff $\Phi_A(\rho) = \Phi_{BA}(\rho)$. In its turn, $\Phi_A(\rho) = \Phi_{BA}(\rho)$ for the following cases:

1. $[A, B] = 0 \quad \forall \rho$, meaning that the operators share the same set of eigenstates;
2. $\rho = \mathbb{1}/d \quad \forall A, B$, meaning the state has no off-diagonal terms, i.e., no quantum probabilities, behaving essentially like a classical state [31];
3. $\rho = B_j$ for A, B forming mutually unbiased bases (MUBs) [53], i.e., $| \langle a_i | b_j \rangle |^2 = \frac{1}{d}$.

These are the cases for compatible contexts according to this definition, which we will also refer to, from now on, as Martins-Savi-Angelo (MSA) definition.

In the search for a context that reduces MSA incompatibility definition to a measurement incompatibility, the authors found that if $\mathbb{C} = \{A_k, A, B\}$, with ρ being an eigenstate of A, then the quantifier would lose its dependence on the initial ρ, with compatibility depending only on the choice of measurements. Ergo, measurement compatibility is contained in this definition.

3.2 CRITICISM

Following the publication introducing the MSA context incompatibility definition, some relevant criticism arose, namely Ref. [24]. The main criticism of this reference is that the definition does not consider generalized measurements. They especially criticized this as a restriction to the third agent, Eve, who, in the condition of a spy, should be described as generally as possible. This is, of course, fair. Mainly because the MSA definition cannot be trivially generalized for POVMs, requiring, as they have shown, the introduction of instruments – which brings its own array of problems to the definition, including the loss of compatibility claim for essentially classic systems.

Hence, they formulated a generalization of the MSA definition as follows: now Alice performs a POVM, $\mathbb{A}$, on a quantum state $\rho \in \mathcal{T}(\mathcal{H})$ using the $\mathbb{A}$-compatible instrument $\mathcal{I}'_{\mathbb{A}} = \{\Phi_{\mathbb{A},x}\}$, such that $\Lambda'_{\mathbb{A}} = \sum_x \Phi_{\mathbb{A},x}$, generating the ensemble $\mathcal{E} = \{p_x, \rho_x\}$, where $p_x = \mathrm{Tr}[\Phi_{\mathbb{A},x}(\rho)]$ and $\rho_x = \frac{\Phi_{\mathbb{A},x}(\rho)}{p_x}$. This ensemble has information content given by the Holevo bound [54]:

$$\chi(\rho, \mathcal{I}'_{\mathbb{A}}) = S(\Lambda'_{\mathbb{A}}(\rho)) - \sum_x p_x S(\rho_x). \tag{3.5}$$

Similarly, Eve performs $\mathcal{I}'_{\mathbb{B}} = \{\Phi_{\mathbb{B},y}\}$, such that $\Lambda'_{\mathbb{B}} = \sum_y \Phi_{\mathbb{B},y}$, generating $\mathcal{E}_{\mathbb{B}} = \{p_x, \Lambda'_{\mathbb{B}}(\rho_x)\}$. The information remaining afterward is then

$$\chi(\rho, \mathcal{I}'_{\mathbb{A}}, \mathcal{I}'_{\mathbb{B}}) = S(\Lambda'_{\mathbb{B}}(\Lambda'_{\mathbb{A}}(\rho))) - \sum_x p_x S(\Lambda'_{\mathbb{B}}(\rho_x)). \tag{3.6}$$

Bob, expecting information $\chi(\rho, \mathcal{I}'_{\mathbb{A}})$, now receives $\chi(\rho, \mathcal{I}'_{\mathbb{A}}, \mathcal{I}'_{\mathbb{B}})$. The information leakage is then measured as

$$\begin{aligned} I_C^H(\rho,\mathcal{I}'_{\mathbb{A}},\mathcal{I}'_{\mathbb{B}}) &= \chi(\rho,\mathcal{I}'_{\mathbb{A}}) - \chi(\rho,\mathcal{I}'_{\mathbb{A}},\mathcal{I}'_{\mathbb{B}}) \\ &= S(\Lambda'_{\mathbb{A}}(\rho)) - S(\Lambda'_{\mathbb{B}}(\Lambda'_{\mathbb{A}}(\rho))) + \sum_x p_x S(\Lambda'_{\mathbb{B}}(\rho_x)) - \sum_x p_x S(\rho_x). \end{aligned} \tag{3.7}$$

Considering that Alice, trying to use a communication channel, would give preference to ensembles with maximally accessible information, a.k.a., maximum $\chi(\rho,\mathcal{I}'_{\mathbb{A}})$; and that Eve, as a spy, would try to get as much information with as little leakage as possible to avoid detection, a.k.a., minimum $I_C^H(\rho,\mathcal{I}'_{\mathbb{A}},\mathcal{I}'_{\mathbb{B}})$, the definition is thus given:

Definition 3.2. *Context incompatibility is the resource encoded in a context $\mathbb{C} = \{\rho,\mathbb{A},\mathbb{B}\}$ that allows one to test the safety of the channel against information leakage. This resource is quantified via $\mathfrak{I}(\mathbb{C}) = I_C^H(\rho,\mathcal{I}_{\mathbb{A},max},\mathcal{I}_{\mathbb{B}})$, where $\mathcal{I}_{\mathbb{A},max}$ is the $\mathbb{A}$-compatible instrument that maximizes $\chi(\rho,\mathcal{I}_{\mathbb{A}})$. Operationally, it is the proper information leakage in the channel caused by an external measurement on the state.*

For this definition, compatible contexts are the ones for $\mathfrak{I}(\mathbb{C}) = 0$, which happens whenever the observables commute, but not for maximally mixed states $\rho = \mathbb{1}/d$. This differs in a significant way from the MSA definition, especially argument-wise: the basis for proposing an incompatibility of physical contexts was precisely that for essentially classical states, like $\rho = \mathbb{1}/d$, there should not be incompatibility.

Furthermore, an additional point of critique worth highlighting is that the MSA definition and this generalized measurement definition are both proposed entirely in a quantum mechanical framework. This is not necessarily a problem, but it is another restriction implied in the definitions.

3.3 TAKEAWAY MESSAGE

Some important remarks on this chapter are:

- The MSA definition is innovative in the manner it included the state in the century-old conversation around incompatibility;
- The MSA definition of context incompatibility is based on a leakage protocol, which has the advantage of allowing an easy translation into resource theory, but the disadvantage of further constraining the framework;
- The definition succeeds in including expected cases for compatibility: commutativity (which it is also shown that their quantifier reduces to for specific choices of state), and the essentially classically behaving state;
- It is not free of criticism, especially in its lack of generality.

Part III

The Proposal

4 THEORY-INDEPENDENT CONTEXT INCOMPATIBILITY PROPOSAL

In the previous chapter, we argued the importance of including the state in the definition of incompatibility based on how it vanishes for classical states. Moreover, the definitions presented thus far have all been strictly quantum – quantum states, quantum observables, quantum entropies, etc. In forethought of post-quantum theories, like the ones hinted to exist by the open questions of quantum mechanics and by the realization there exists correlations that respect relativity theory (in the form of the no-signaling theorem) but are not constrained to quantum theory and still violate Bell-type inequalities [26], produce steering [27] and could be used to generate better than quantum key distribution protocols for cryptography tasks [28], we were driven to search for a context incompatibility definition that is based on more fundamental elements, entirely independent of any theory's particular formalism.

In order to do this, we concluded that basing the definition on probabilities was reasonable, as it also could be applied to the theory we know to present incompatibility – quantum mechanics. Furthermore, with the focus on a probabilistic approach, it is sensible to not restrict ourselves to a single realization with a well-defined output, consequently leading to the use of non-selective measurements, where the exact output is not specified.

On that account, this Part is reserved for presenting the proposal of a theory-independent context incompatibility definition, a case study, quantifiers, and further explorations.

4.1 DEFINITION

There is a concept of probability in many physical theories, with quantum mechanics itself being regarded by some as strictly probabilistic. As such, an attempt at a completely general incompatibility statement based on probabilities seems appropriate. In order to propose the aforementioned statement, we also consider that it is possible to define non-selective measurements in any theory. In quantum mechanics, for example, they are given by (1.30), but a definition for a generic theory could be achieved through the idea of "omitting" or "forgetting" the result of a measurement. In terms of probabilities, it could look something like measuring the generic measurable physical quantity $\mathscr{Y} = \{y_j\}$ in a state $\mathscr{E}$, which characterizes the preparation of the system, with a probability of obtaining y_j given by $p_{\mathscr{E}}(y_j)$. If a measurement of another generic measurable physical quantity $\mathscr{X} = \{x_i\}$ had been done in this system, with outcome x_i, the probability of obtaining y_i would now be written as $p_{\mathscr{E}}(y_j|x_i)$. The omission or forgetting the result of the measurement of $\mathscr{X}$ can be accounted for in this system through the multiplication of $p_{\mathscr{E}}(y_j|x_i)$ by the probability of obtaining the specific x_i in the first place, $p_{\mathscr{E}}(x_i)$, and summing for all the possible outcomes of $\mathscr{X}$, as Equation (4.1) shows:

$$p_{\mathscr{E}}(y_j) \xrightarrow{x_i} p_{\mathscr{E}}(y_j|x_i) \xrightarrow{\text{omission}} \sum_i^d p_{\mathscr{E}}(y_j|x_i)p_{\mathscr{E}}(x_i) =: p_{\mathscr{M}_{\mathscr{X}}(\mathscr{E})}(y_j). \tag{4.1}$$

We then call $\mathscr{M}_{\mathscr{X}}(\mathscr{E})$ the non-selective measurement map of $\mathscr{X}$ onto $\mathscr{E}$.

Notice that if this was a quantum mechanical context, $\mathbb{C} = \{\rho, A, B\}$, simply substituting the probabilities given by (1.17) with the quantum mechanical non-selective map (1.30) and using the cyclic and linear property of the trace would give

$$\begin{aligned} p_{\Phi_A(\rho)}(b_j) &= \mathrm{Tr}[B_j\Phi_A(\rho)] = \mathrm{Tr}[B_j\sum_i^d p_\rho(a_i)A_i] = \sum_i^d p_\rho(a_i)\underbrace{\mathrm{Tr}[B_jA_i]}_{p(b_j|a_i)} \\ &= \sum_i^d p(b_j|a_i)p_\rho(a_i), \end{aligned} \tag{4.2}$$

which indicates that (4.1) is a good way to generalize a definition for the probability distribution of a physical quantity on a system that underwent a non-selective measurement map.

With these elements, we propose that it is reasonable to expect that if a context is compatible, having performed a measurement without extracting information about the result from the system should not change the probability distribution of another measurement. Simply put, we define context compatibility as follows.

Definition 4.1. *If the non-selective measurement of $\mathscr{Y}$ does not alter the probability distribution of $\mathscr{X}$, and vice versa, for a given preparation $\mathscr{E}$, that is*

$$p_{\mathscr{E}}(x_i) = p_{\mathscr{M}_{\mathscr{Y}}(\mathscr{E})}(x_i), \tag{4.3a}$$

$$p_{\mathscr{E}}(y_j) = p_{\mathscr{M}_{\mathscr{X}}(\mathscr{E})}(y_j), \tag{4.3b}$$

then the context $\mathbb{C} = \{\mathscr{E}, \mathscr{X}, \mathscr{Y}\}$ is said to be compatible.

This definition is, indeed, entirely independent of theory inasmuch as it does not require the specification of any details about interactions, dynamics, and algebraic structure. On that account, we can check now how it behaves for a classical system in order to determine if it predicts compatibility, as one would expect.

For this, consider the classical statistical mechanical theory for a one-dimensional particle, described in the phase space through the coordinates related to position and momentum, (q, p), and through the probability density $\varrho_t(q, p)$, solution to the Liouville equation [55] $\partial_t\varrho_t = \{H, \varrho_t\}$. After measuring the position, $\mathscr{Q}$, and obtaining, say, $\bar{q}$ as a result, we describe the new (normalized) probability density through a Dirac delta as follows:

$$\varrho_t(q,p) \xrightarrow{\bar{q}} \frac{\delta(q-\bar{q})\varrho_t(\bar{q},p)}{\int \mathrm{d}q'\mathrm{d}p'\,\delta(q'-\bar{q})\varrho_t(\bar{q},p')} = \frac{\delta(q-\bar{q})\varrho_t(\bar{q},p)}{\int \mathrm{d}p'\,\varrho_t(\bar{q},p')} =: \varrho_t(q,p|\bar{q}), \tag{4.4}$$

where the denominator corresponds to the probability density $\varrho_t(\bar{q})$ of obtaining $\bar{q}$ when measuring $\mathscr{Q}$.

Now, to describe a non-selective map in this scenario we resort to the previously described process of omission, where we multiply (4.4) by the probability density of obtaining $\mathscr{Q}$, $\varrho_t(\bar{q})$, and sum (or, in this case, integrate) over all possible outcomes:

$$\begin{aligned}\mathscr{M}_{\mathscr{Q}}\left(\varrho_t(q,p)\right) &:= \int \mathrm{d}\bar{q}\,\varrho_t(q,p|\bar{q})\varrho_t(\bar{q}) = \int \mathrm{d}\bar{q}\,\frac{\delta(q-\bar{q})\varrho_t(\bar{q},p)}{\int \mathrm{d}p'\,\varrho_t(\bar{q},p')} \int \mathrm{d}p''\,\varrho_t(\bar{q},p'') \\ &= \varrho_t(q,p). \end{aligned} \tag{4.5}$$

The formulation for the measurement of $\mathscr{P}$ is equivalent and would lead to the same conclusion – applying a non-selective map on this classical description does not change the probability density, and, according to Definition (4.1), this means there is *always* compatibility for such a classical statistical distribution.

As for a quantum mechanical context $\mathbb{C} = \{\rho, A, B\}$, using the quantum probabilities alongside the non-selective map description, allows us to rewrite Equations (4.3) from Definition 4.1 as

$$\mathrm{Tr}\left\{A_i\left[\rho - \Phi_B(\rho)\right]\right\} = 0 \quad \text{or} \quad \mathrm{Tr}\left\{\left[A_i - \Phi_B(A_i)\right]\rho\right\} = 0, \tag{4.6a}$$

$$\mathrm{Tr}\left\{B_j\left[\rho - \Phi_A(\rho)\right]\right\} = 0 \quad \text{or} \quad \mathrm{Tr}\left\{\left[B_j - \Phi_A(B_j)\right]\rho\right\} = 0, \tag{4.6b}$$

where the relation $\mathrm{Tr}[A_i\Phi_B(\rho)] = \mathrm{Tr}[A_i\sum_j^d B_j\rho B_j] = \mathrm{Tr}[\sum_j^d B_jA_iB_j\rho] = \mathrm{Tr}[\Phi_B(A_i)\rho]$ was used to obtain the extra two equations (Heisenberg picture [56]), indicating that for a context to be compatible it must satisfy one of the (4.6a) equations <u>and</u> one of the (4.6b) equations simultaneously. It is possible to show that this will only occur if $\phi_A(\rho) = \phi_{AB}(\rho)$ and $\phi_B(\rho) = \phi_{BA}(\rho)$.

Theorem 4.2. $\mathrm{Tr}[A_i\rho] = \mathrm{Tr}[A_i\phi_B(\rho)]$ *iff* $\phi_A(\rho) = \phi_{AB}(\rho)$*, where* $A = \sum_i^d a_iA_i$ *and* $B = \sum_l^d b_lB_l$ *are PVMs.*

Proof. Take $\mathrm{Tr}[A_i\rho] = \mathrm{Tr}[A_i\phi_B(\rho)]$ and multiply both sides of it by the identity relation $\sum_i^d A_i = \mathbb{1}$, obtaining $\sum_i^d \mathrm{Tr}[A_i\rho]A_i = \sum_i^d \mathrm{Tr}[A_i\phi_B(\rho)]A_i$. From the definition of the non-selective map, we recognize that this is equal to $\phi_A(\rho) = \phi_{AB}(\rho)$. On the other hand, if $\phi_A(\rho) = \phi_{AB}(\rho)$, we can write it explicitly and multiply both sides by one of the projectors of A, say A_k, obtaining $\sum_i^d \mathrm{Tr}[A_i\rho]A_iA_k = \sum_i^d \mathrm{Tr}[A_i\phi_B(\rho)]A_iA_k$. Since $A_iA_k = \delta_{ik}A_i$, this is equal to $\mathrm{Tr}[A_i\rho]A_i = \mathrm{Tr}[A_i\phi_B(\rho)]A_i$. Then, taking the trace of both sides knowing that $\mathrm{Tr}[A_i] = 1$, we finally obtain $\mathrm{Tr}[A_i\rho] = \mathrm{Tr}[A_i\phi_B(\rho)]$, which completes the proof. ■

Theorem (4.2) is, of course, also valid for $\mathrm{Tr}[B_j\rho] = \mathrm{Tr}[B_j\Phi_A(\rho)]$, and so, in the quantum framework, we can write (4.6) equivalently as

$$\Phi_A(\rho) = \Phi_{AB}(\rho), \tag{4.7a}$$
$$\Phi_B(\rho) = \Phi_{BA}(\rho). \tag{4.7b}$$

Observe that $\Phi_A(\rho) = \Phi_{AB}(\rho)$ is curiously similar to what appeared in the MSA definition, $\Phi_A(\rho) = \Phi_{BA}(\rho)$, even though the path for arriving at (4.7a) is completely different. Regardless of their similarity, the equations are nonetheless different and we cannot apply the same steps in order to find the specific cases that satisfy (4.7) exactly as it is done for the MSA definition. For the purpose of investigating which cases satisfy compatibility for quantum contexts, we turn to the Bloch representation since it provides an easier visualization of results.

4.1.1 Case Study

For a two-dimensional qubit acting on $\mathcal{H} \simeq \mathbb{C}^2$, we know from Subsection 1.1.6 that the density operator is given by $\rho_r = (1/2)[\mathbb{1} + \mathbf{r}\cdot\sigma]$ and the observables projectors are given by $A_i = (1/2)[\mathbb{1} + \mathbf{a}_i\cdot\sigma]$ and $B_j = (1/2)[\mathbb{1} + \mathbf{b}_j\cdot\sigma]$. The trace property of the Pauli matrices $\mathrm{Tr}[\sigma] = 0$ and the multiplication property $(\mathbf{m}\cdot\sigma)(\mathbf{n}\cdot\sigma) = (\mathbf{m}\cdot\mathbf{n})\mathbb{1} + \imath(\mathbf{m}\times\mathbf{n})\cdot\sigma$ allow us to then arrive at the expressions for the probabilities:

$$\begin{aligned}
p_\rho(a_i) = \mathrm{Tr}[A_i\rho] &= \mathrm{Tr}\left[\frac{1}{4}[\mathbb{1} + \mathbf{r}\cdot\sigma][\mathbb{1} + \mathbf{a}_i\cdot\sigma]\right] \\
&= \mathrm{Tr}\left[\frac{1}{4}[\mathbb{1} + \mathbb{1}\mathbf{a}_i\cdot\sigma + \mathbb{1}\mathbf{r}\cdot\sigma + \underbrace{(\mathbf{r}\cdot\sigma)(\mathbf{a}_i\cdot\sigma)}_{(\mathbf{r}\cdot\mathbf{a}_i)\mathbb{1}+\imath(\mathbf{r}\times\mathbf{a}_i)\cdot\sigma}]\right] \\
&= \frac{1}{4}[2 + 2(\mathbf{r}\cdot\mathbf{a}_i)] = \frac{1}{2}[1 + (\mathbf{r}\cdot\mathbf{a}_i)], \qquad (4.8a) \\
p(a_i|b_j) = \mathrm{Tr}[A_iB_j] &= \mathrm{Tr}\left[\frac{1}{4}[\mathbb{1} + \mathbf{a}_i\cdot\sigma][\mathbb{1} + \mathbf{b}_j\cdot\sigma]\right] \\
&= \mathrm{Tr}\left[\frac{1}{4}[\mathbb{1} + \mathbb{1}\mathbf{a}_i\cdot\sigma + \mathbb{1}\mathbf{b}_j\cdot\sigma + (\mathbf{a}_i\cdot\sigma)(\mathbf{b}_j\cdot\sigma)]\right] \\
&= \frac{1}{2}[1 + (\mathbf{a}_i\cdot\mathbf{b}_j)], \qquad (4.8b)
\end{aligned}$$

and similarly, $p_\rho(b_j) = \frac{1}{2}[1 + (\mathbf{r}\cdot\mathbf{b}_j)]$.

With these probabilities and considering that we can write the projectors vectors as $\mathbf{a}_i = (-1)^i\hat{\mathbf{a}}$ and $\mathbf{b}_j = (-1)^j\hat{\mathbf{b}}$, we can easily write the action of the maps as:

$$\phi_A(\rho) = \sum_i^2 p_\rho(a_i)A_i = \sum_i^2 \frac{1}{4}[1 + (\mathbf{r}\cdot\mathbf{a}_i)][\mathbb{1} + (\mathbf{a}_i\cdot\sigma)]$$

$$
\begin{aligned}
&= \sum_{i}^{2}\frac{1}{4}[\underbrace{\mathbb{1}}_{2\mathbb{1}} + (\mathbf{a}_i\cdot\sigma) + \mathbb{1}(\mathbf{r}\cdot\mathbf{a}_i) + \underbrace{(\mathbf{r}\cdot\mathbf{a}_i)(\mathbf{a}_i\cdot\sigma)}_{(-1)^{2i}(\mathbf{r}\cdot\hat{\mathbf{a}})(\hat{\mathbf{a}}\cdot\sigma)}] \\
&= \frac{1}{2}[\mathbb{1} + (\mathbf{r}\cdot\hat{\mathbf{a}})(\hat{\mathbf{a}}\cdot\sigma)], \qquad (4.9a)
\end{aligned}
$$

$$
\begin{aligned}
\phi_{AB}(\rho) &= \sum_{i,j}^{2} p_\rho(b_j)p(a_i|b_j)A_i = \sum_{i,j}^{2}\frac{1}{8}[1+(\mathbf{r}\cdot\mathbf{b}_j)][1+(\mathbf{a}_i\cdot\mathbf{b}_j)][\mathbb{1}+(\mathbf{a}_i\cdot\sigma)] \\
&= \sum_{i,j}^{d}\frac{1}{8}[\underbrace{\mathbb{1}}_{4\mathbb{1}} + \mathbb{1}(\mathbf{a}_i\cdot\mathbf{b}_j) + \mathbb{1}(\mathbf{r}\cdot\mathbf{b}_j) + \mathbb{1}(\mathbf{r}\cdot\mathbf{b}_j)(\mathbf{a}_i\cdot\mathbf{b}_j) + (\mathbf{a}_i\cdot\sigma) \\
&\quad + (\mathbf{a}_i\cdot\mathbf{b}_j)(\mathbf{a}_i\cdot\sigma) + (\mathbf{r}\cdot\mathbf{b}_j)(\mathbf{a}_i\cdot\sigma) + \underbrace{(\mathbf{r}\cdot\mathbf{b}_j)(\mathbf{a}_i\cdot\mathbf{b}_j)(\mathbf{a}_i\cdot\sigma)}_{(-1)^{2i+2j}(\mathbf{r}\cdot\hat{\mathbf{b}})(\hat{\mathbf{a}}\cdot\hat{\mathbf{b}})(\hat{\mathbf{a}}\cdot\sigma)}] \\
&= \frac{1}{2}[\mathbb{1} + (\mathbf{r}\cdot\hat{\mathbf{b}})(\hat{\mathbf{a}}\cdot\hat{\mathbf{b}})(\hat{\mathbf{a}}\cdot\sigma)]. \qquad (4.9b)
\end{aligned}
$$

Correspondingly, $\phi_B(\rho) = \frac{1}{2}[\mathbb{1} + (\mathbf{r}\cdot\hat{\mathbf{b}})(\hat{\mathbf{b}}\cdot\sigma)]$ and $\phi_{BA}(\rho) = \frac{1}{2}[\mathbb{1} + (\mathbf{r}\cdot\hat{\mathbf{a}})(\hat{\mathbf{b}}\cdot\hat{\mathbf{a}})(\hat{\mathbf{b}}\cdot\sigma)]$.

Therefore, the criteria (4.7) for a context compatibility is met when

$$
\begin{aligned}
\phi_A(\rho) = \phi_{AB}(\rho) \implies \frac{1}{2}[\mathbb{1} + (\mathbf{r}\cdot\hat{\mathbf{a}})(\hat{\mathbf{a}}\cdot\sigma)] &= \frac{1}{2}[\mathbb{1} + (\mathbf{r}\cdot\hat{\mathbf{b}})(\hat{\mathbf{a}}\cdot\hat{\mathbf{b}})(\hat{\mathbf{a}}\cdot\sigma)] \\
(\mathbf{r}\cdot\hat{\mathbf{a}})(\hat{\mathbf{a}}\cdot\sigma) &= (\mathbf{r}\cdot\hat{\mathbf{b}})(\hat{\mathbf{a}}\cdot\hat{\mathbf{b}})(\hat{\mathbf{a}}\cdot\sigma) \\
(\mathbf{r}\cdot\hat{\mathbf{a}}) &= (\mathbf{r}\cdot\hat{\mathbf{b}})(\hat{\mathbf{a}}\cdot\hat{\mathbf{b}}) \\
\mathbf{r}\cdot[\hat{\mathbf{a}} - \hat{\mathbf{b}}(\hat{\mathbf{a}}\cdot\hat{\mathbf{b}})] &= 0, \qquad (4.10a) \\
\phi_B(\rho) = \phi_{BA}(\rho) \implies (\mathbf{r}\cdot\hat{\mathbf{b}})(\hat{\mathbf{b}}\cdot\sigma) &= (\mathbf{r}\cdot\hat{\mathbf{a}})(\hat{\mathbf{b}}\cdot\hat{\mathbf{a}})(\hat{\mathbf{b}}\cdot\sigma) \\
(\mathbf{r}\cdot\hat{\mathbf{b}}) &= (\mathbf{r}\cdot\hat{\mathbf{a}})(\hat{\mathbf{b}}\cdot\hat{\mathbf{a}}) \\
\mathbf{r}\cdot[\hat{\mathbf{b}} - \hat{\mathbf{a}}(\hat{\mathbf{b}}\cdot\hat{\mathbf{a}})] &= 0. \qquad (4.10b)
\end{aligned}
$$

Rewriting Equations (4.10), we get

$$
\begin{aligned}
\mathbf{r}\cdot\underbrace{[\hat{\mathbf{a}} - \hat{\mathbf{b}}(\hat{\mathbf{a}}\cdot\hat{\mathbf{b}})]}_{\gamma} &= 0 \implies \|\mathbf{r}\|\|\gamma\|\cos\theta_{\mathbf{r},\gamma} = 0, \qquad (4.11a) \\
\mathbf{r}\cdot\underbrace{[\hat{\mathbf{b}} - \hat{\mathbf{a}}(\hat{\mathbf{b}}\cdot\hat{\mathbf{a}})]}_{\eta} &= 0 \implies \|\mathbf{r}\|\|\eta\|\cos\theta_{\mathbf{r},\eta} = 0, \qquad (4.11b)
\end{aligned}
$$

where $\theta_{\mathbf{r},\gamma}$ and $\theta_{\mathbf{r},\eta}$ are the angles between the $\mathbf{r}$ and γ, and the $\mathbf{r}$ and η vectors, respectively. From these relations it is straightforward to distinguish tree cases for compatibility:

1. $\|\mathbf{r}\| = 0$, which is the maximally mixed density operator $\rho = \mathbb{1}/d$. This translates to commutativity of ρ with both A and B, as $[A,\rho] = \imath(\hat{\mathbf{a}}\times\mathbf{r})\cdot\sigma = 0$ and equivalently for $[B,\rho]$.

   This is a reasonable result, as density operators of this form are free of any quantum features (off-diagonal terms are zero in all bases), having only classical probabilities and behaving essentially as a classical state would –therefore, no incompatibility is expected;

2. $\|\gamma\| = 0$, which happens only when the observables vectors are parallel, $(\hat{\mathbf{a}} \cdot \hat{\mathbf{b}}) = 1$. In this case, $\|\boldsymbol{\eta}\|$ is also automatically zero, ensuring that both (4.11) equations are simultaneously satisfied. This translates to commuting observables, since $[A, B] = (\hat{\mathbf{a}} \cdot \sigma)(\hat{\mathbf{b}} \cdot \sigma) - (\hat{\mathbf{b}} \cdot \sigma)(\hat{\mathbf{a}} \cdot \sigma) = 2\imath(\hat{\mathbf{a}} \times \hat{\mathbf{b}}) \cdot \sigma = 0$.

   Again, this is a desired result, as it shows that this definition encompasses commutativity. In addition, along with 1, this result shows that our notion of context compatibility is readily satisfied by the commutativity of any two operators of the context under scrutiny. In other words, pairwise noncommutativity is necessary for context incompatibility;

3. $\cos\theta_{\mathbf{r},\gamma} = 0$ and $\cos\theta_{\mathbf{r},\eta} = 0$. To scrutinize this case, we consider the form of the vector $\gamma = [\hat{\mathbf{a}} - \hat{\mathbf{b}}(\hat{\mathbf{a}} \cdot \hat{\mathbf{b}})]$: $(\hat{\mathbf{a}} \cdot \hat{\mathbf{b}})$ is just a number making $\hat{\mathbf{b}}$ smaller, and the subtraction of $\hat{\mathbf{a}}$ and this smaller vector puts γ perpendicular to $\hat{\mathbf{b}}$. The same can be said about $\boldsymbol{\eta}$, making it perpendicular to $\hat{\mathbf{a}}$. As cosines are zero for perpendicular vectors, this means that for both of those terms to be zero simultaneously, $\mathbf{r}$ must be perpendicular to the plane formed by $\hat{\mathbf{a}}$ <u>and</u> $\hat{\mathbf{b}}$. This translates to the commutativity relation $[\rho, [A, B]] = -2(\mathbf{r} \times (\hat{\mathbf{a}} \times \hat{\mathbf{b}})) \cdot \sigma = 0$.

   This gives an entire class of contexts that are compatible and elude the expected cases 1 and 2.

Because of its similarity in form, testing if the MSA definition also encompassed this seemingly new case for compatibility could be interesting. But since this case does not appear as effortlessly for the MSA definition, we needed to build a quantifier for our definition to then be able to compare them.

4.2 QUANTIFICATION OF CONTEXT INCOMPATIBILITY

Besides comparing with previous notions, knowing *how much* incompatibility a context possesses is interesting since this proposed definition of context incompatibility may be explored in the future in a resource theory frame, in the same way as the MSA incompatibility and the joint measurability have been [57]. To quantify this incompatibility, then, we tried a similar approach as the one presented in Part II through the entropy, not only to compare with the MSA incompatibility but also having in sight the connection of entropy with information and how it is commonly used to interpret incompatibility.

Considering that the relative entropy between two quantum density operators ρ and σ (1.28) is equal to zero *iff* $\rho = \sigma$, the relative entropy seems like a prime candidate for a quantifier, as, from the Definition (4.7), there is no incompatibility whenever $\Phi_A(\rho) = \phi_{AB}(\rho)$ and $\Phi_B(\rho) = \phi_{BA}(\rho)$.

However, because the definition is complementary (as in, there is an equation for $\Phi_A(\rho)$ and an equation for $\Phi_B(\rho)$ that have to be simultaneously satisfied), we consider that a quantifier must contain both equations to truly represent the status of incompatibility of the context, and not allow for a claim of compatibility when only one of the equations is zero. Hence, we arrived at the proposed quantifier for the quantum case:

$$\frac{S(\Phi_A(\rho)\|\Phi_{AB}(\rho)) + S(\Phi_B(\rho)\|\Phi_{BA}(\rho))}{2} =: \mathcal{I}_{\mathbb{C}}. \tag{4.12}$$

This proposal of a relative entropy-based quantifier was also inspired by the existence of the analogous probabilistic version given by the Kullback-Liebler divergence (1.41) that also has the property of only being zero if the things being compared are equal. This allows us to write a quantifier for the general proposal (4.3) similarly:

$$\frac{D(p_{\mathscr{E}}(x_i)\|p_{\mathscr{M}_{\mathscr{Y}}(\mathscr{E})}(x_i)) + D(p_{\mathscr{E}}(y_j)\|p_{\mathscr{M}_{\mathscr{X}}(\mathscr{E})}(y_j))}{2} =: \mathcal{D}_{\mathbb{C}}. \tag{4.13}$$

It is possible to show that Equations (4.12) and (4.13) are equal for quantum contexts:

Proof. Start from $S(\Phi_A(\rho)\|\Phi_{AB}(\rho)) = \mathrm{Tr}[\Phi_A(\rho)\log\Phi_A(\rho)] - \mathrm{Tr}[\Phi_A(\rho)\log\Phi_{AB}(\rho)]$ and substitute the maps,

$$\begin{aligned}
S(\Phi_A(\rho)\|\Phi_{AB}(\rho)) &= \mathrm{Tr}\Big[\sum_i^d p_\rho(a_i)A_i\Big(\log\sum_{i'}^d p_\rho(a_{i'})A_{i'} - \log\sum_{i'}^d p_{\Phi_B(\rho)}(a_{i'})A_{i'}\Big)\Big] \\
&= \sum_i^d p_\rho(a_i)\log p_\rho(a_i) - \sum_i^d p_\rho(a_i)\log p_{\Phi_B(\rho)}(a_i) \\
&= D(p_\rho(a_i)\|p_{\Phi_B(\rho)}(a_i)),
\end{aligned}$$

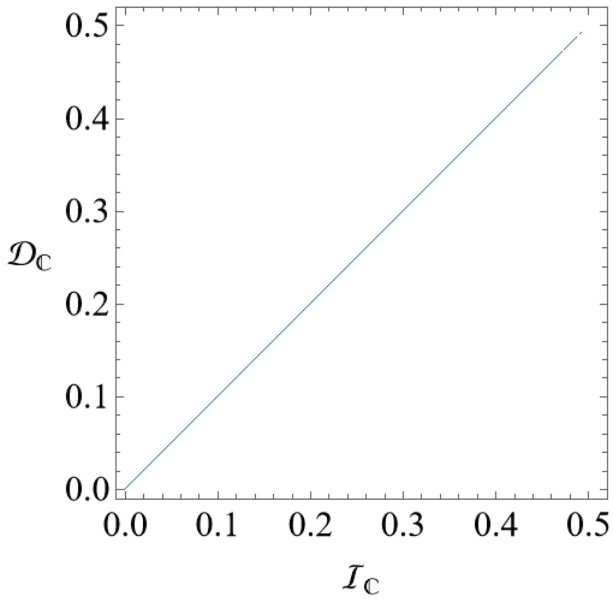
where the property $f(A)\,|a\rangle = f(a)\,|a\rangle$, with $f(\cdot)$ being any function, A an Hermitian operator with orthonormal base $|a\rangle$ and eigenvalue a, was used. ■

As such, this form of quantifier suits our much sought-after independence of theory aspect. The interchangeability is illustrated in Figure 4, made for a two-dimensional system (qubit) in the Bloch representation, as done in Subsection 4.1.1, with 10^6 randomized sets of contexts $\{\mathbf{r}, \hat{\mathbf{a}}, \hat{\mathbf{b}}\}$ with logarithm base 2 for the entropy.

One of the goals of this work was to compare the predictions of our definition to the MSA ones, however, since the MSA definition is not complementary (it has a statement only regarding the Φ_A map and not the Φ_B one) and thus clearly the proposed definition would be stronger, we opted, for fairness' sake, to add the complementary term to its quantifier, turning it into

$$\frac{S(\Phi_A(\rho)\|\Phi_{BA}(\rho)) + S(\Phi_B(\rho)\|\Phi_{AB}(\rho))}{2} =: \mathcal{I}_{\mathbb{C}}^{MSA}, \tag{4.14}$$

Figure 4 – Relationship between $\mathcal{I}_{\mathbb{C}}$, Equation (4.12), and $\mathcal{D}_{\mathbb{C}}$, Equation (4.13), for the qubit in the Bloch representation.



where it is possible, as already mentioned previously, to write the relative entropy in this case as $S(\rho\|\Phi_A(\rho)) = S(\Phi_A(\rho)) - S(\rho)$ and therefore

$$\mathcal{I}_{\mathbb{C}}^{MSA} = \frac{S(\Phi_{BA}(\rho)) - S(\Phi_A(\rho)) + S(\Phi_{AB}(\rho)) - S(\Phi_B(\rho))}{2}. \tag{4.15}$$

It is noteworthy that, unlike what happened for our definition, the quantification of the MSA incompatibility using the Kullback-Leibler divergence is not equal to the one based on the relative entropy,

$$\frac{D(p_\rho(a_i)\|p_{\Phi_{BA}(\rho)}(a_i)) + D(p_\rho(b_j)\|p_{\Phi_{AB}(\rho)}(b_j))}{2} =: \mathcal{D}_{\mathbb{C}}^{MSA}, \tag{4.16}$$

as the Kullback-Leibler divergence poses only as a lower bound.

Proof. Start from $S(\Phi_A(\rho)\|\Phi_{BA}(\rho)) = \mathrm{Tr}[\Phi_A(\rho)\log\Phi_A(\rho)] - \mathrm{Tr}[\Phi_A(\rho)\log\Phi_{BA}(\rho)]$ and write explicitly the map $\Phi_A(\rho)$, applying the trace:

$$\begin{aligned} S(\Phi_A(\rho)\|\Phi_{BA}(\rho)) &= \sum_i^d p_\rho(a_i)\log p_\rho(a_i) - \mathrm{Tr}\left[\sum_i^d p_\rho(a_i)A_i\log\Phi_{BA}(\rho)\right] \\ &= \sum_i^d p_\rho(a_i)\log p_\rho(a_i) - \sum_i^d p_\rho(a_i)\langle a_i|\log\Phi_{BA}(\rho)|a_i\rangle. \end{aligned}$$

Then, add and subtract $\sum_i^d p_\rho(a_i)\log\langle a_i|\Phi_{BA}(\rho)|a_i\rangle = \sum_i^d p_\rho(a_i)\log p_{\Phi_{BA}(\rho)}(a_i)$ to obtain

$$S(\Phi_A(\rho)\|\Phi_{BA}(\rho)) = \sum_i^d p_\rho(a_i)\log p_\rho(a_i) - \sum_i^d p_\rho(a_i)\langle a_i|\log\Phi_{BA}(\rho)|a_i\rangle$$

$$+ \sum_i^d p_\rho(a_i) \log \langle a_i | \Phi_{BA}(\rho) | a_i \rangle - \sum_i^d p_\rho(a_i) \log \langle a_i | \Phi_{BA}(\rho) | a_i \rangle ,$$

where one can recognize the Kullback-Liebler divergence so that we have

$$\begin{aligned} S(\Phi_A(\rho) \| \Phi_{BA}(\rho)) = D(p_\rho(a_i) \| p_{\Phi_{BA}(\rho)}(a_i)) & - \sum_i^d p_\rho(a_i) \langle a_i | \log \Phi_{BA}(\rho) | a_i \rangle \\ & + \sum_i^d p_\rho(a_i) \log \langle a_i | \Phi_{BA}(\rho) | a_i \rangle \\ = D(p_\rho(a_i) \| p_{\Phi_{BA}(\rho)}(a_i) & + \sum_i^d p_\rho(a_i) \Big[\langle a_i | - \log \Phi_{BA}(\rho) | a_i \rangle \\ & -(- \log \langle a_i | \Phi_{BA}(\rho) | a_i \rangle) \Big]. \end{aligned}$$

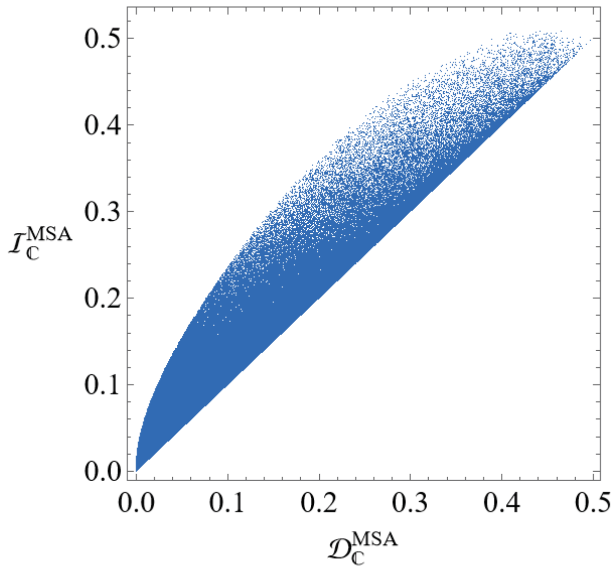
In this form, one can easily use Jensen's inequality [58] to state that $\langle a_i | - \log \Phi_{BA}(\rho) | a_i \rangle \geq - \log \langle a_i | \Phi_{BA}(\rho) | a_i \rangle$ and therefore

$$\begin{aligned} S(\Phi_A(\rho) \| \Phi_{BA}(\rho)) - D(p_\rho(a_i) \| p_{\Phi_{BA}(\rho)}(a_i)) \geq 0 \\ S(\Phi_A(\rho) \| \Phi_{BA}(\rho)) \geq D(p_\rho(a_i) \| p_{\Phi_{BA}(\rho)}(a_i)), \end{aligned}$$

ergo, $\mathcal{I}_{\mathbb{C}}^{MSA} \geq \mathcal{D}_{\mathbb{C}}^{MSA}$, thus ending the proof. ■

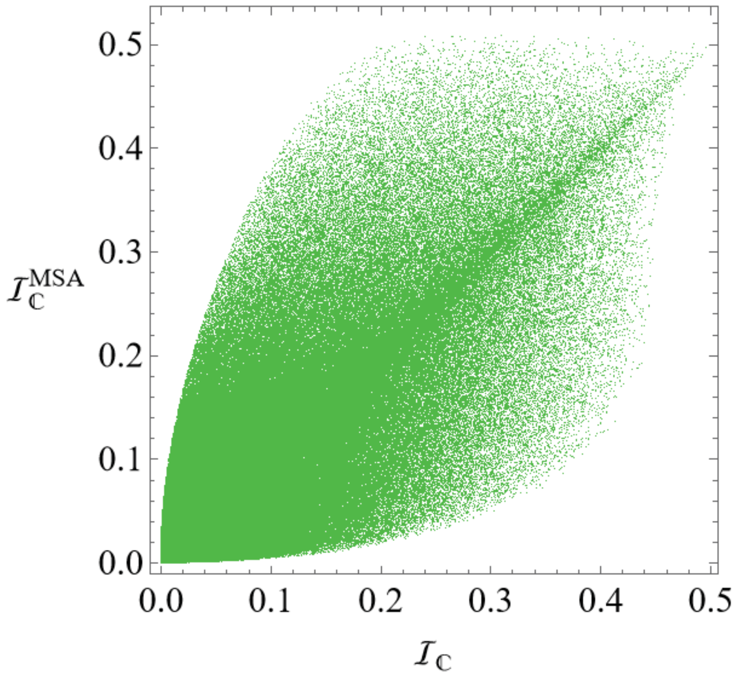
This bound can be verified in Figure 5, also built for the qubit in the Bloch representation with 10^6 points of randomized contexts $\mathbb{C} = \{\mathbf{r}, \hat{\mathbf{a}}, \hat{\mathbf{b}}\}$. This in itself depicts some differences between our proposal and the existing MSA definition, where the first has the advantage of allowing a probabilistic approach interchangeably.

Figure 5 – Relationship between the complementary quantum quantifier version for the MSA incompatibility $\mathcal{I}_{\mathbb{C}}^{MSA}$, given by (4.15), and the attempt for a Kullback-Liebler divergence version $\mathcal{D}_{\mathbb{C}}^{MSA}$, given by (4.16). Note how the latter is indeed a lower bound.



Despite this distinction, numerical verification for the qubit in the Bloch scenario shows that quantifiers $\mathcal{I}_{\mathbb{C}}$ and $\mathcal{I}_{\mathbb{C}}^{MSA}$ agree, for over a million randomly generated contexts, whether the context is compatible. That is, there was no occurrence of $\mathcal{I}_{\mathbb{C}} = 0$ when $\mathcal{I}_{\mathbb{C}}^{MSA} \neq 0$, and vice versa. The behavior of the quantifiers for the same randomly generated contexts can be seen in Figure 6, created with the same constraints as the others, showcasing that they detect compatibility equally, but may quantify incompatibility differently, which indicates that the differences of the quantifiers are limited to how much incompatibility there is in the incompatible contexts, and not on the central issue of there being compatibility or not. This type of graph forming a "leaf" shape is not unheard of for the comparison of different quantifiers and poses no direct problem [59].

Figure 6 – Relationship between $\mathcal{I}_{\mathbb{C}}$, given by the Equation (4.12), and $\mathcal{I}_{\mathbb{C}}^{MSA}$, given by the Equation (4.15), showing that they are equivalent, albeit not equal.



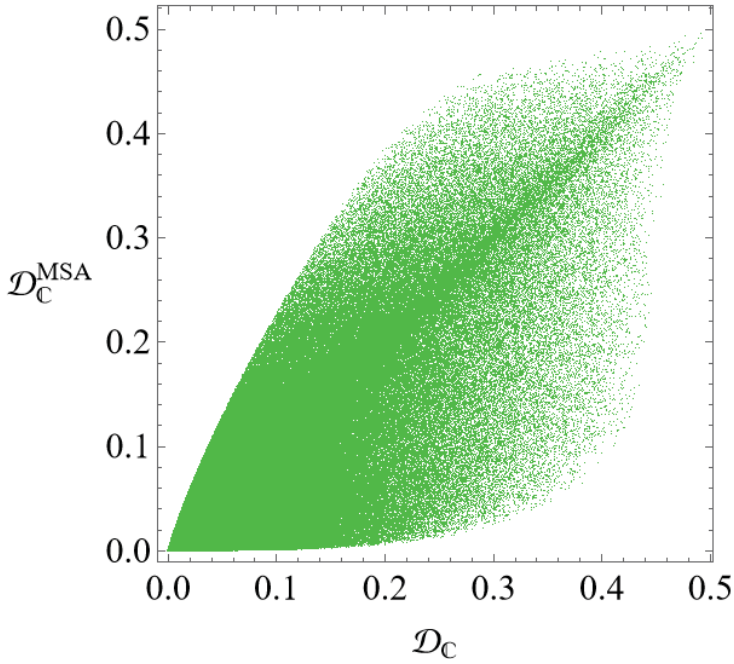
On closer examination of Figure 6, it is possible to see a prominent line across the diagonal where $\mathcal{I}_{\mathbb{C}} = \mathcal{I}_{\mathbb{C}}^{MSA}$. This line is given, in the Bloch representation following Subsection 4.1.1, by the contexts for which

$$\left[\sum_{\epsilon}^{d} \frac{1+\epsilon|\mathbf{r}\cdot\hat{\mathbf{a}}|}{2} \log \frac{1+\epsilon|(\mathbf{r}\cdot\hat{\mathbf{b}})(\hat{\mathbf{a}}\cdot\hat{\mathbf{b}})|}{2}\right] + \left[\sum_{\epsilon}^{d} \frac{1+\epsilon|\mathbf{r}\cdot\hat{\mathbf{b}}|}{2} \log \frac{1+\epsilon|(\mathbf{r}\cdot\hat{\mathbf{a}})(\hat{\mathbf{a}}\cdot\hat{\mathbf{b}})|}{2}\right] \tag{4.17}$$
$$= \left[\sum_{\epsilon}^{d} \frac{1+\epsilon|(\mathbf{r}\cdot\hat{\mathbf{a}})(\hat{\mathbf{a}}\cdot\hat{\mathbf{b}})|}{2} \log \frac{1+\epsilon|(\mathbf{r}\cdot\hat{\mathbf{a}})(\hat{\mathbf{a}}\cdot\hat{\mathbf{b}})|}{2}\right] + \left[\sum_{\epsilon}^{d} \frac{1+\epsilon|(\mathbf{r}\cdot\hat{\mathbf{b}})(\hat{\mathbf{a}}\cdot\hat{\mathbf{b}})|}{2} \log \frac{1+\epsilon|(\mathbf{r}\cdot\hat{\mathbf{b}})(\hat{\mathbf{a}}\cdot\hat{\mathbf{b}})|}{2}\right],$$

where $\epsilon = \pm 1$. This equality holds, distinguishably, if i) $(\hat{\mathbf{a}} \cdot \hat{\mathbf{b}}) = \pm 1$ and ii) $(\hat{\mathbf{a}} \cdot \hat{\mathbf{b}}) = 0$. The first case is simply the commutativity $[A, B] = 0$, with both quantifiers equal to zero and therefore predicting compatibility. The second case is for perpendicular observables (MUBs), or maximally incompatible observables. In this case, the choice of the state vector $\mathbf{r}$ just interferes with how much incompatibility there is in the system, but both quantifiers give out the same amount.

Recollecting the discussion on the Kullback-Leibler divergence versions of the quantifiers, Figure 7 comparing $\mathcal{D}_{\mathbb{C}}^{MSA}$ and $\mathcal{D}_{\mathbb{C}}$, made with the same constraints as the others, clearly showcases the relations $\mathcal{D}_{\mathbb{C}} = \mathcal{I}_{\mathbb{C}}$ and $\mathcal{D}_{\mathbb{C}}^{MSA} \leq \mathcal{I}_{\mathbb{C}}^{MSA}$. For the divergences, we have also found numerically that $\mathcal{D}_{\mathbb{C}}^{MSA}$ quantifies compatibility for contexts that $\mathcal{D}_{\mathbb{C}}$ also quantifies compatibility. This suggests that, although the quantifiers based on entropy and on divergence are not equivalent for the MSA definition, they seem to agree on which contexts are compatible.

Figure 7 – Relationship between $\mathcal{D}_{\mathbb{C}}$, given by (4.13), in the x-axis, and $\mathcal{D}_{\mathbb{C}}^{MSA}$, given by (4.16), in the y-axis, accentuating both the inequality $\mathcal{D}_{\mathbb{C}}^{MSA} \leq \mathcal{I}_{\mathbb{C}}^{MSA}$ and the equality $\mathcal{D}_{\mathbb{C}} = \mathcal{I}_{\mathbb{C}}$ when compared to Figure 6.



We can conclude that, when the complementary expression is added to the MSA definition, both definitions will yield the same result regarding the compatibility of the quantum context and are valuable in their own sense, the choice between them is thus more dependent on the objective of the implementer: the MSA definition brings along a leakage detection protocol and has a ready-to-use formulation in resource theory, as well as a simpler quantifier in the quantum framework; whereas our model is based on concepts that any probabilistic physical theory has, allowing for an easy adaptation for any description of probability and non-selective measurements (for example, generalizing the quantum measurements to POVMs), also providing an interchangeable probability-based quantifier, which can be useful in laboratory implementations.

4.3 RELATION TO MEASUREMENT INCOMPATIBILITY

When the concept of context incompatibility is introduced as a generalization, it is natural then to ask if and when it reduces to an incompatibility related to the measurements alone. As an educated guess, based on what was found for the MSA context incompatibility and that it would be reasonable to expect that for states compatible to one of the observables, incompatibility should only come from the second measurement choice, we tested the context in the form $\mathbb{C} = \{A_k, A, B\}$, where $\rho = A_k$ is an eigenstate of the observable $A = \sum_i^d a_i A_i$.

For such context, the maps required for the quantifier $\mathcal{I}_{\mathbb{C}}$ are:

$$\begin{aligned}
\Phi_A(\rho) &= \sum_i^d A_i \rho A_i = \sum_i^d |\langle a_i|a_k\rangle|^2 A_i = A_k, && \text{(4.18a)}\\
\Phi_B(\rho) &= \sum_j^d B_j \rho B_j = \sum_j^d |\langle b_j|a_k\rangle|^2 B_j, && \text{(4.18b)}\\
\Phi_{AB}(\rho) &= \Phi_A(\Phi_B(\rho)) = \sum_{ij}^d |\langle b_j|a_k\rangle|^2 |\langle b_j|a_i\rangle|^2 A_i, && \text{(4.18c)}\\
\Phi_{BA}(\rho) &= \Phi_B(\Phi_A(\rho)) = \sum_j^d |\langle b_j|a_k\rangle|^2 B_j = \Phi_B(\rho), && \text{(4.18d)}
\end{aligned}$$

where the eigenvalues for each of the maps are given by

$$\begin{aligned}
\lambda_{\Phi_A(\rho)} &= |\langle a_i|a_k\rangle|^2 = \delta_{ik}, && \text{(4.19a)}\\
\lambda_{\Phi_B(\rho)} &= |\langle b_j|a_k\rangle|^2, && \text{(4.19b)}\\
\lambda_{\Phi_{AB}(\rho)} &= \sum_j^d |\langle b_j|a_k\rangle|^2 |\langle b_j|a_i\rangle|^2, && \text{(4.19c)}\\
\lambda_{\Phi_{BA}(\rho)} &= |\langle b_j|a_k\rangle|^2. && \text{(4.19d)}
\end{aligned}$$

Therefore, we have the relative entropies

$$\begin{aligned}
S(\Phi_A(\rho)\|\Phi_{AB}(\rho)) &= \sum_i^d \delta_{ik}\log\delta_{ik} - \sum_i^d \delta_{ik}\log\Big[\sum_j^d |\langle b_j|a_k\rangle|^2 |\langle b_j|a_i\rangle|^2\Big]\\
&= -\log\Big[\sum_j^d |\langle b_j|a_i\rangle|^4\Big], && \text{(4.20a)}\\
S(\Phi_B(\rho)\|\Phi_{BA}(\rho)) &= 0, && \text{(4.20b)}
\end{aligned}$$

and the quantifier results in

$$\mathcal{I}_{\{A_k,A,B\}} = -\frac{1}{2}\log\Big[\sum_j^d |\langle b_j|a_i\rangle|^4\Big], \tag{4.21}$$

which is, in fact, an incompatibility quantifier of the measurements alone. Note that if A and B form MUB and $|\langle b_j|a_i\rangle|^2 = 1/d$, this yields $\mathcal{I}_{\{A_k,A,B\}} = \frac{1}{2}\log d$, which is the maximum amount found for $\mathcal{I}_{\mathbb{C}}$ numerically, for over a million contexts with $d = 2$ in the Bloch representation.

It follows similarly for $\rho = B_l$, arriving at $\mathcal{I}_{\{B_l,A,B\}} = -\frac{1}{2}\log\left[\sum_i^d |\langle a_i|b_j\rangle|^4\right]$. As such, we can conclude that if the state is an eigenstate of one of the observables, the context incompatibility is reduced to a measurement incompatibility; and if the observables therein form a MUB, then the incompatibility is maximum, as expected.

As a more general test aligned with this, we inputted a context in which the state ρ is A-real, i.e., $\rho = \Phi_A(\sigma) = \sum_k \langle a_k|\,\sigma\,|a_k\rangle\, A_k = \sum_k p_\sigma(a_k) A_k = \sum_k s_k A_k$, forming the context $\mathbb{C} = \{\Phi_A(\sigma), A, B\}$. The non-selective maps for such context have the form:

$$\Phi_A(\rho) = \sum_i A_i\rho A_i = \sum_i A_i \sum_k s_k A_k A_i = \sum_k s_k A_k, \tag{4.22a}$$

$$\Phi_B(\rho) = \sum_j B_j \sum_k s_k A_k B_j = \sum_{jk} s_k|\langle b_j|a_k\rangle|^2 B_j, \tag{4.22b}$$

$$\Phi_{AB}(\rho) = \sum_i A_i \sum_{jk} s_k|\langle b_j|a_k\rangle|^2 B_j A_i = \sum_{ijk} s_k|\langle b_j|a_k\rangle|^2|\langle b_j|a_i\rangle|^2 A_i, \tag{4.22c}$$

$$\Phi_{BA}(\rho) = \sum_j B_j \sum_k s_k A_k B_j = \sum_{jk} s_k|\langle b_j|a_k\rangle|^2 B_j = \Phi(\rho), \tag{4.22d}$$

where the eigenvalues for each of the maps are given by

$$\lambda_{\Phi_A(\rho)} = s_k, \tag{4.23a}$$

$$\lambda_{\Phi_B(\rho)} = \sum_k s_k|\langle b_j|a_k\rangle|^2, \tag{4.23b}$$

$$\lambda_{\Phi_{AB}(\rho)} = \sum_{jk} s_k|\langle b_j|a_k\rangle|^2|\langle b_j|a_i\rangle|^2, \tag{4.23c}$$

$$\lambda_{\Phi_{AB}(\rho)} = \sum_k s_k|\langle b_j|a_k\rangle|^2. \tag{4.23d}$$

Hence, the relative entropies are

$$\begin{aligned} S(\Phi_A(\rho)\|\Phi_{AB}(\rho)) &= \sum_k s_k\log s_k - \sum_{lk}\langle a_l|\, s_k A_k \log\Big[\sum_{i'j'k'} s_{k'}|\langle b_{j'}|a_{k'}\rangle|^2|\langle b_{j'}|a_{i'}\rangle|^2 A_{i'}\Big]\,|a_l\rangle \\ &= \sum_k s_k\log s_k - \sum_{lk}\delta_{lk}s_k\log\Big[\sum_{i'j'k'}\delta_{li'}s_{k'}|\langle b_{j'}|a_{k'}\rangle|^2|\langle b_{j'}|a_{i'}\rangle|^2\Big] \\ &= \sum_k s_k\log s_k - \sum_k s_k\log\Big[\sum_{j'k'} s_{k'}|\langle b_{j'}|a_{k'}\rangle|^2|\langle b_{j'}|a_k\rangle|^2\Big], \end{aligned} \tag{4.24a}$$

$$S(\Phi_B(\rho)\|\Phi_{BA}(\rho)) = 0, \tag{4.24b}$$

and the quantifier yields

$$\mathcal{I}_{\{\Phi_A(\sigma),A,B\}} = \frac{1}{2}\Bigg[\sum_k s_k\log s_k - \sum_k s_k\log\Big[\sum_{j'k'} s_{k'}|\langle b_{j'}|a_{k'}\rangle|^2|\langle b_{j'}|a_k\rangle|^2\Big]\Bigg]. \tag{4.25}$$

This equation has a lasting dependence on s_k, which is in turn dependent on the initial state of the system. Therefore, this shows that having states of A-reality (equivalently, B-reality) is not enough to reduce the context incompatibility to a measurement-only incompatibility.

The tests in this Subsection have also been performed in the Bloch representation, which allows for an easier illustration of the conclusions, and are presented in Appendix 1.

4.4 TAKEAWAY MESSAGE

The most important points of this chapter are:

- A theory-independent context incompatibility definition proposal was achieved;
- The new definition has tight connections with the MSA definition, but they are not identical and were not reached through the same premisses;
- When equipped with a corresponding complementary equation, the MSA definition detects compatibility for the same two-dimensional PVM-based contexts as the proposed definition, but still quantifies the amount of incompatibility in the incompatible contexts differently;
- The proposed definition offers an interchangeable probability-based quantifier;
- Even though most of the results showcased are restricted to projective measurements (and some to two dimensions), the definition being theory-independent and having a probability-based quantifier leaves room for generalizations.

CONCLUDING REMARKS

The article on the MSA context incompatibility definition [23] brought a novel approach in its inclusion of the state to the century-old conversation around incompatibility. Remarkably, it included the expected cases for compatibility, commuting observables and essentially classically behaving states, and reduced to a measurement incompatibility for well-chosen contexts, solidifying the context-based definition as a strong generalized approach to incompatibility.

Nonetheless, it was clear from the beginning that it had strong constraints, be it on the strict use of PVM observables, or the entirely quantum framework. The first constrain was tackled by the authors in [24], but the generalization of measurements made it so that the definition no longer assessed compatibility for maximally mixed states, which was a key point of the argument for a context incompatibility in the first place. The latter constraint was the focus of the work presented in this dissertation, motivated by the idea that generalizing the definition of context incompatibility to more fundamental frameworks could allow it to thrive in post-quantum theories, which have been sought after through the stipulation of more basic postulates to quantum mechanics with the hope that this would reproduce its accomplishments while resolving some of its open questions.

Thus, we proposed a theory-independent context incompatibility for arbitrary probabilistic theories and, as such, we accomplished the *raison d'être* of allowing the verification of a context's incompatibility regardless of its quantum status.

When compared to the MSA definition, we found validation for what our proposed definition quantified, as they agree regarding which contexts are compatible. Yet, we see that the definitions are not exactly the same, quantifying the amount of incompatibility differently and each of them has pros and cons in its applications and frameworks.

The definition proposed in this dissertation shows great potential. In future work, we intend to focus on the generalization of the case study to an arbitrary dimension, as well as expanding the definition to multipartite systems. Interestingly, the multipartite approach opens the possibility of exploring the connections of this definition with other uncanny quantum effects such as Bell non-locality and steering, which have previously been tightly linked to some incompatibility notions [18, 19]. There is room for the future exploration of generalized measurements in this definition as well since they also have probabilistic formulations.

BIBLIOGRAPHY

1 HEISENBERG, W. Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik. **Zeitschrift fur Physik**, v. 43, n. 3-4, p. 172–198, Mar. 1927. DOI: 10.1007/BF01397280. Cit. on pp. 10, 28, 35.

2 BOHR, N. The Quantum Postulate and the Recent Development of Atomic Theory1. **Nature**, v. 121, p. 580–590, Apr. 1928. DOI: 10.1038/121580a0. Cit. on p. 10.

3 EINSTEIN, A.; PODOLSKY, B.; ROSEN, N. Can Quantum-Mechanical Description of Physical Reality Be Considered Complete? **Phys. Rev.**, American Physical Society, v. 47, p. 777–780, 10 May 1935. DOI: 10.1103/PhysRev.47.777. Cit. on p. 10.

4 FINE, A. Hidden Variables, Joint Probability, and the Bell Inequalities. **Phys. Rev. Lett.**, American Physical Society, v. 48, p. 291295, 5 Feb. 1982. DOI: 10.1103/PhysRevLett.48.291. Cit. on p. 10.

5 MASANES, L.; ACIN, A.; GISIN, N. General properties of nonsignaling theories. **Phys. Rev. A**, American Physical Society, v. 73, p. 012112, 1 Jan. 2006. DOI: 10.1103/PhysRevA.73.012112. Cit. on p. 10.

6 WOLF, M. M.; PEREZ-GARCIA, D.; FERNANDEZ, C. Measurements Incompatible in Quantum Theory Cannot Be Measured Jointly in Any Other No-Signaling Theory. **Phys. Rev. Lett.**, American Physical Society, v. 103, p. 230402, 23 Dec. 2009. DOI: 10.1103/PhysRevLett.103.230402. Cit. on p. 10.

7 HENSEN, B. e. a. Loophole-free Bell inequality violation using electron spins separated by 1.3 kilometres. **Nature**, v. 526, p. 682–686, 2015. DOI: 10.1038/nature15759. Cit. on pp. 10, 33.

8 THE Nobel Prize in Physics 2022. [S.l.: s.n.]. https://www.nobelprize.org/prizes/physics/2022/summary/. Cit. on p. 10.

9 ROBERTSON, H. P. The Uncertainty Principle. **Phys. Rev.**, American Physical Society, v. 34, p. 163–164, 1 July 1929. DOI: 10.1103/PhysRev.34.163. Cit. on pp. 10, 28.

10 MANDL, F. **Quantum Mechanics**. England: Wiley, 1992. Cit. on p. 10.

11 GRIFFITHS, D. J.; SCHROETER, D. F. **Introduction to Quantum Mechanics**. 3. ed. [S.l.]: Cambridge University Press, 2018. DOI: 10.1017/9781316995433. Cit. on p. 10.

12 MÜLLER-KIRSTEN, H. J. W. **Introduction to Quantum Mechanics**. 2nd. [S.l.]: WORLD SCIENTIFIC, 2012. DOI: 10.1142/8428. Cit. on p. 10.

13 WERNER, R. F.; FARRELLY, T. Uncertainty from Heisenberg to Today. **Foundations of Physics**, v. 49, p. 460–491, 6 June 2019. DOI: 10.1007/s10701-019-00265-z. Cit. on pp. 10, 29.

14 BUSCH, P.; HEINONEN, T.; LAHTI, P. Heisenberg's uncertainty principle. **Physics Reports**, v. 452, n. 6, p. 155–176, 2007. ISSN 0370-1573. DOI: https://doi.org/10.1016/j.physrep.2007.05.006. Cit. on pp. 10, 29, 31.

15 BUSCH, P. Indeterminacy Relations and Simultaneous Measurements in Quantum Theory. **International Journal of Theoretical Physics**, v. 24, n. 1, p. 63–92, Jan. 1985. DOI: 10.1007/BF00670074. Cit. on pp. 11, 30.

16 ALI, S. T.; CARMELI, C.; HEINOSAARI, T.; TOIGO, A. Commutative POVMs and Fuzzy Observables. **Foundations of Physics**, v. 39, p. 593–612, 6 June 2009. DOI: 10.1007/s10701-009-9292-y. Cit. on p. 11.

17 BUSCH, P. Unsharp reality and joint measurements for spin observables. **Phys. Rev. D**, American Physical Society, v. 33, p. 2253–2261, 8 Apr. 1986. DOI: 10.1103/PhysRevD.33.2253. Cit. on p. 11.

18 HEINOSAARI, T.; MIYADERA, T.; ZIMAN, M. An invitation to quantum incompatibility. **Journal of Physics A: Mathematical and Theoretical**, IOP Publishing, v. 49, n. 12, p. 123001, Feb. 2016. DOI: 10.1088/1751-8113/49/12/123001. Cit. on pp. 11, 54.

19 UOLA, R.; COSTA, A. C. S.; NGUYEN, H. C.; GÜHNE, O. Quantum steering. **Rev. Mod. Phys.**, American Physical Society, v. 92, p. 015001, 1 Mar. 2020. DOI: 10.1103/RevModPhys.92.015001. Cit. on pp. 11, 54.

20 UOLA, R.; MORODER, T.; GÜHNE, O. Joint Measurability of Generalized Measurements Implies Classicality. **Phys. Rev. Lett.**, American Physical Society, v. 113, 16z403, 16 Oct. 2014. DOI: 10.1103/PhysRevLett.113.160403. Cit. on p. 11.

21 QUINTINO, M. T.; VÉRTESI, T.; BRUNNER, N. Joint Measurability, Einstein-Podolsky-Rosen Steering, and Bell Nonlocality. **Phys. Rev. Lett.**, American Physical Society, v. 113, p. 160402, 16 Oct. 2014. DOI: 10.1103/PhysRevLett.113.160402. Cit. on p. 11.

22 HEINOSAARI, T.; WOLF, M. M. Nondisturbing quantum measurements. **Journal of Mathematical Physics**, v. 51, n. 9, p. 092201, 2010. DOI: 10.1063/1.3480658. Cit. on pp. 11, 31.

23 MARTINS, E.; SAVI, M. F.; ANGELO, R. M. Quantum incompatibility of a physical context. **Phys. Rev. A**, American Physical Society, v. 102, p. 050201, 5 Nov. 2020. DOI: 10.1103/PhysRevA.102.050201. Cit. on pp. 11, 35, 36, 54.

24 MITRA, A.; SHARMA, G.; GHOSH, S. Information leak and incompatibility of physical context: A modified approach. **Phys. Rev. A**, American Physical Society, v. 104, p. 032225, 3 Sept. 2021. DOI: 10.1103/PhysRevA.104.032225. Cit. on pp. 12, 37, 54.

25 BOHM, D. A Suggested Interpretation of the Quantum Theory in Terms of "Hidden" Variables. I. **Phys. Rev.**, American Physical Society, v. 85, p. 166–179, 2 Jan. 1952. DOI: 10.1103/PhysRev.85.166. Cit. on p. 12.

26 POPESCU, S.; ROHRLICH, D. Quantum nonlocality as an axiom. **Foundations of Physics**, v. 24, p. 379–385, 3 Mar. 1994. DOI: 10.1007/BF02058098. Cit. on pp. 12, 40.

27 CAVALCANTI, P. J.; SELBY, J. H.; SIKORA, J.; GALLEY, T. D.; SAINZ, A. B. Post-quantum steering is a stronger-than-quantum resource for information processing. **npj Quantum Information**, v. 8, 1 June 2022. DOI: 10.1038/s41534-022-00574-8. Cit. on pp. 12, 40.

28 BARRETT, J.; HARDY, L.; KENT, A. No Signaling and Quantum Key Distribution. **Phys. Rev. Lett.**, American Physical Society, v. 95, p. 010503, 1 June 2005. DOI:

10.1103/PhysRevLett.95.010503. Available from: https://link.aps.org/doi/10.1103/PhysRevLett.95.010503. Cit. on pp. 12, 40.

29 NIELSEN, M. A.; CHUANG, I. L. **Quantum Computation and Quantum Information**. 10th ed. USA: Cambridge University Press, 2011. ISBN 1107002176. Cit. on pp. 13, 23, 26, 29.

30 COHEN-TANNOUDJI, C.; DIU, B.; LALOE, F. **Quantum mechanics**. New York: Wiley, 1978. v. 1. Cit. on pp. 13, 27, 28.

31 LANDI, G. T. **Quantum Information and Quantum Noise**. São Paulo: University of São Paulo, July 2018. Available from: http://www.fmt.if.usp.br/~gtlandi/lecture-notes-2.pdf. Cit. on pp. 13, 37.

32 MARTINS, E. **Quantifying Quantum Incompatibility**. 2020. MA thesis – UFPR, Curitiba, PR. Cit. on pp. 13, 21, 23, 36.

33 SAVI, M. F. **Quantum Resource Covariance: Correlations, Coherence, and Incompatibility**. 2021. PhD thesis – UFPR, Curitiba, PR. Cit. on pp. 13, 30, 33.

34 JUNIOR, A. C. O. **Non-Classicalities in Quantum Walks and an Axiomatic Approach to Quantum Realism**. 2022. PhD thesis – UFPR, Curitiba, PR. Cit. on pp. 13, 22.

35 M. A. NEUMARK. Extremal spectral functions of a symmetric operator. **Izv. Akad. Nauk SSSR Ser. Mat.**, p. 327–344, 4 1947. Available from: http://mi.mathnet.ru/im3002. Cit. on pp. 17, 35.

36 NEUMANN, J. v. Wahrscheinlichkeitstheoretischer Aufbau der Quantenmechanik. **Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse**, v. 1927, p. 245–272, 1927. Available from: http://eudml.org/doc/59230. Cit. on p. 18.

37 NEUMANN, J. v. **Mathematical Foundations of Quantum Mechanics: New Edition**. Ed. by Nicholas A. Wheeler. Trans. by Robert T. Beyer. United States: Princeton University Press, 2018. Cit. on p. 18.

38 GRIFFITHS, R. B. **Quantum Channels, Kraus Operators, POVMs**. [S.l.]: Quantum Computation and Quantum Information Theory Course, Spring 2014. Available from: https://quantum.phys.cmu.edu/QCQI/qitd412.pdf. Cit. on p. 20.

39 BILOBRAN, A. L. O.; ANGELO, R. M. A measure of physical reality. **Europhysics Letters**, EDP Sciences, IOP Publishing and Società Italiana di Fisica, v. 112, n. 4, p. 40005, Nov. 2015. DOI: 10.1209/0295-5075/112/40005. Cit. on p. 21.

40 ROSS, S. M. **A First Course in Probability**. 5th ed. Upper Saddle River, N.J.: Prentice Hall, 1998. Available from: http://julio.staff.ipb.ac.id/files/2015/02/Ross_8th_ed_English.pdf. Cit. on pp. 23, 24.

41 STIRZAKER, D.; GRIMMETT, G. **Probability and Random Processes**. 2nd ed. United Kingdom: Clarendon Press, 1992. Cit. on pp. 23, 24.

42 KULLBACK, S.; LEIBLER, R. A. On Information and Sufficiency. **The Annals of Mathematical Statistics**, Institute of Mathematical Statistics, v. 22, n. 1, p. 79–86, 1951. DOI: 10.1214/aoms/1177729694. Cit. on p. 24.

43 SHANNON, C. E. A mathematical theory of communication. **The Bell System Technical Journal**, v. 27, n. 3, p. 379–423, 1948. DOI: 10.1002/j.1538-7305.1948.tb01338.x. Cit. on p. 24.

44 SAKURAI, J. J. **Modern quantum mechanics**. Reading, MA: Addison-Wesley, 1994. Available from: https://cds.cern.ch/record/1167961. Cit. on pp. 26, 29.

45 CAUCHY, A.-L. Note II. Sur les formules qui résultent de l'emploi du signe $>$ ou $<$, et sur les moyennes entre plusieurs quantités. In: COURS d'analyse de l'École Royale Polytechnique, Première Partie: Analyse algébrique. Paris: Debure frères, 1821. Cit. on p. 28.

46 SCHWARZ, H. A. Ueber ein die Flächen kleinsten Flächeninhalts betreffendes Problem der Variationsrechnung. In: GESAMMELTE Mathematische Abhandlungen: Erster Band. Berlin, Heidelberg: Springer Berlin Heidelberg, 1890. P. 223–269. DOI: 10.1007/978-3-642-50665-9_11. Cit. on p. 28.

47 SHANKAR, R. The Heisenberg Uncertainty Relations. In: PRINCIPLES of Quantum Mechanics. New York, NY: Springer US, 1994. P. 237–246. ISBN 978-1-4757-0576-8. DOI: 10.1007/978-1-4757-0576-8_9. Cit. on p. 29.

48 KOK, P. Uncertainty Relations. In: A First Introduction to Quantum Physics. Cham: Springer International Publishing, 2018. P. 171–194. ISBN 978-3-319-92207-2. DOI: 10.1007/978-3-319-92207-2_9. Cit. on p. 29.

49 HEINOSAARI, T.; REITZNER, D.; STANO, P. Notes on Joint Measurability of Quantum Observables. **Foundations of Physics**, v. 38, p. 1133–1147, 12 Dec. 2008. DOI: 10.1007/s10701-008-9256-7. Cit. on p. 30.

50 LÜDERS, G. Concerning the state-change due to the measurement process. **Annalen der Physik**, v. 518, n. 9, p. 663–670, Sept. 2006. DOI: 10.1002/andp.2006518090410.1002/andp.200610207. Cit. on p. 32.

51 PERES, A. (Ed.). Information and Thermodynamics. In: **Quantum Theory: Concepts and Methods**. Dordrecht: Springer Netherlands, 2002. P. 260–297. ISBN 978-0-306-47120-9. DOI: 10.1007/0-306-47120-5_9. Cit. on p. 35.

52 STECK, D. A. **Measurement and POVMs**. Willamette: Recent Developments in Quantum Mechanics and Quantum Information Lecture Notes, Spring 2009. Available from: http://atomoptics-nas.uoregon.edu/~dsteck/teaching/09spring/phys610/notes/povms.pdf. Cit. on p. 35.

53 DURT, T.; ENGLERT, B.-G.; BENGTSSON, I.; ŻYCZKOWSKI, K. On Mutually Unbiased Bases. **International Journal of Quantum Information**, v. 08, n. 04, p. 535–640, 2010. DOI: 10.1142/S0219749910006502. Cit. on p. 37.

54 HOLEVO, A. S. Statistical problems in quantum physics. In________. **Proceedings of the Second Japan-USSR Symposium on Probability Theory**. Berlin, Heidelberg: Springer Berlin Heidelberg, 1973. P. 104–119. ISBN 978-3-540-46956-8. Cit. on p. 37.

55 LIOUVILLE, J. Note sur la Théorie de la Variation des constantes arbitraires. fre. **Journal de Mathématiques Pures et Appliquées**, p. 342–349, 1838. Available from: http://eudml.org/doc/234417. Cit. on p. 41.

56 KUKIN, V. D. Heisenberg Representation. In: **Encyclopaedia of Mathematics**. Organized by: M. Hazewinkel. Boston, MA: Springer US, 1995. v. 3, p. 4. DOI: 10.1007/978-1-4899-3793-3. Cit. on p. 42.

57 HEINOSAARI, T.; KIUKAS, J.; REITZNER, D. Noise robustness of the incompatibility of quantum measurements. **Phys. Rev. A**, American Physical Society, v. 92, p. 022115, 2 Aug. 2015. DOI: 10.1103/PhysRevA.92.022115. Cit. on p. 45.

58 JENSEN, J. L. W. V. Sur les fonctions convexes et les inégalités entre les valeurs moyennes. **Acta Mathematica**, Institut Mittag-Leffler, v. 30, none, p. 175–193, 1906. DOI: 10.1007/BF02418571. Cit. on p. 48.

59 COSTA, A.; BEIMS, M.; ANGELO, R. Generalized discord, entanglement, Einstein–Podolsky–Rosen steering, and Bell nonlocality in two-qubit systems under (non-)Markovian channels: Hierarchy of quantum resources and chronology of deaths and births. **Physica A: Statistical Mechanics and its Applications**, v. 461, p. 469–479, 2016. ISSN 0378-4371. DOI: https://doi.org/10.1016/j.physa.2016.05.068. Cit. on p. 49.

Appendix

APPENDIX 1 – BLOCH REPRESENTATION VERSION

The following results were obtained for the two-dimensional Bloch representation as seen in Subsection 1.1.6.

For contexts $\mathbb{C} = \{A_k, A, B\}$, we can write the state vector $\mathbf{r}$ as $\mathbf{r} = \mathbf{a_k} = (-1)^k\hat{\mathbf{a}}$, so that the maps (4.9) become

$$\Phi_A(\rho) = \frac{1}{2}[\mathbb{1} + (\mathbf{r}\cdot\hat{\mathbf{a}})(\hat{\mathbf{a}}\cdot\sigma)] = \frac{1}{2}[\mathbb{1} + (-1)^k(\hat{\mathbf{a}}\cdot\sigma)] = A_k, \tag{1.1a}$$
$$\Phi_B(\rho) = \frac{1}{2}[\mathbb{1} + (\mathbf{r}\cdot\hat{\mathbf{b}})(\hat{\mathbf{b}}\cdot\sigma)] = \frac{1}{2}[\mathbb{1} + (-1)^k(\hat{\mathbf{a}}\cdot\hat{\mathbf{b}})(\hat{\mathbf{b}}\cdot\sigma)], \tag{1.1b}$$
$$\Phi_{AB}(\rho) = \frac{1}{2}[\mathbb{1} + (\mathbf{r}\cdot\hat{\mathbf{b}})(\hat{\mathbf{a}}\cdot\hat{\mathbf{b}})(\hat{\mathbf{a}}\cdot\sigma)] = \frac{1}{2}[\mathbb{1} + (-1)^k(\hat{\mathbf{a}}\cdot\hat{\mathbf{b}})^2(\hat{\mathbf{a}}\cdot\sigma)], \tag{1.1c}$$
$$\Phi_{BA}(\rho) = \frac{1}{2}[\mathbb{1} + (\mathbf{r}\cdot\hat{\mathbf{a}})(\hat{\mathbf{b}}\cdot\hat{\mathbf{a}})(\hat{\mathbf{b}}\cdot\sigma)] = \frac{1}{2}[\mathbb{1} + (-1)^k(\hat{\mathbf{a}}\cdot\hat{\mathbf{b}})(\hat{\mathbf{b}}\cdot\sigma)] = \Phi_B(\rho), \tag{1.1d}$$

with respective eigenvalues

$$\lambda^{\epsilon}_{\Phi_A(\rho)} = \frac{1+\epsilon\|(-1)^k\hat{\mathbf{a}}\|}{2} = \frac{1+\epsilon 1}{2} \qquad \epsilon = \pm 1, \tag{1.2a}$$
$$\lambda^{\epsilon}_{\Phi_B(\rho)} = \frac{1+\epsilon\|(-1)^k(\hat{\mathbf{a}}\cdot\hat{\mathbf{b}})\hat{\mathbf{b}}\|}{2} = \frac{1+\epsilon|(\hat{\mathbf{a}}\cdot\hat{\mathbf{b}})|}{2} = \lambda^{\epsilon}_{\Phi_{BA}(\rho)} \qquad \epsilon = \pm 1, \tag{1.2b}$$
$$\lambda^{\epsilon}_{\Phi_{AB}(\rho)} = \frac{1+\epsilon\|(-1)^k(\hat{\mathbf{a}}\cdot\hat{\mathbf{b}})^2\hat{\mathbf{a}}\|}{2} = \frac{1+\epsilon|(\hat{\mathbf{a}}\cdot\hat{\mathbf{b}})^2|}{2} \qquad \epsilon = \pm 1. \tag{1.2c}$$

The entropy of the map $\Phi_A(\rho)$ is $S(\Phi_A(\rho)) = -1\log 1 - 0\log 0 = 0$, therefore, the relative entropies are

$$\begin{aligned} S(\Phi_A(\rho)|\Phi_{AB}(\rho)) =& -S(\Phi_A) - \sum_{\epsilon}\lambda^{\epsilon}_{\Phi_A(\rho)}\log\lambda^{\epsilon}_{\Phi_{AB}(\rho)} \\ =& -1\log\frac{1+|(\hat{\mathbf{a}}\cdot\hat{\mathbf{b}})^2|}{2} - 0\log\frac{1-|(\hat{\mathbf{a}}\cdot\hat{\mathbf{b}})^2|}{2} \\ =& -\log\frac{1+|(\hat{\mathbf{a}}\cdot\hat{\mathbf{b}})^2|}{2}, \end{aligned} \tag{1.3a}$$
$$S(\Phi_B(\rho)|\Phi_{BA}(\rho)) = 0, \tag{1.3b}$$

hence, the quantifier yields

$$\mathcal{I}_{\{A_k,A,B\}} = \frac{S(\Phi_A(\rho)\|\Phi_{AB}(\rho)) + S(\Phi_B(\rho)\|\Phi_{BA}(\rho))}{2} = -\frac{1}{2}\log\frac{1+|(\hat{\mathbf{a}}\cdot\hat{\mathbf{b}})^2|}{2}. \tag{1.4}$$

It is now even clearer than from (4.21) that in the contexts where ρ is an eigenstate of one of the observables the quantifier reduces to a relation between the observables only. This conclusion is not restricted to the analysis of the quantifier, as applying the definition's criteria (4.7) directly with the Equations (1.1) would also entail a measurement incompatibility.

For contexts $\mathbb{C} = \{\Phi_A(\sigma), A, B\}$, we write the state as

$$\rho = \Phi_A(\sigma) = \frac{1}{2}[\mathbb{1} + (\boldsymbol{s}\cdot\hat{\mathbf{a}})(\hat{\mathbf{a}}\cdot\boldsymbol{\sigma})],$$

where s refers to the vector that characterizes the state σ. Therefore, the maps (4.9) become

$$\Phi_A(\rho) = \frac{1}{2}[\mathbb{1} + (\boldsymbol{s}\cdot\hat{\mathbf{a}})(\hat{\mathbf{a}}\cdot\hat{\mathbf{a}})(\hat{\mathbf{a}}\cdot\boldsymbol{\sigma})] = \frac{1}{2}[\mathbb{1} + (\boldsymbol{s}\cdot\hat{\mathbf{a}})(\hat{\mathbf{a}}\cdot\boldsymbol{\sigma})] = \rho, \tag{1.5a}$$

$$\Phi_B(\rho) = \frac{1}{2}[\mathbb{1} + (\boldsymbol{s}\cdot\hat{\mathbf{a}})(\hat{\mathbf{a}}\cdot\hat{\mathbf{b}})(\hat{\mathbf{b}}\cdot\boldsymbol{\sigma})], \tag{1.5b}$$

$$\Phi_{AB}(\rho) = \frac{1}{2}[\mathbb{1} + (\boldsymbol{s}\cdot\hat{\mathbf{a}})(\hat{\mathbf{a}}\cdot\hat{\mathbf{b}})(\hat{\mathbf{a}}\cdot\hat{\mathbf{b}})(\hat{\mathbf{a}}\cdot\boldsymbol{\sigma})] = \frac{1}{2}[\mathbb{1} + (\boldsymbol{s}\cdot\hat{\mathbf{a}})(\hat{\mathbf{a}}\cdot\hat{\mathbf{b}})^2(\hat{\mathbf{a}}\cdot\boldsymbol{\sigma})], \tag{1.5c}$$

$$\Phi_{BA}(\rho) = \frac{1}{2}[\mathbb{1} + (\boldsymbol{s}\cdot\hat{\mathbf{a}})(\hat{\mathbf{a}}\cdot\hat{\mathbf{a}})(\hat{\mathbf{b}}\cdot\hat{\mathbf{a}})(\hat{\mathbf{b}}\cdot\boldsymbol{\sigma})] = \frac{1}{2}[\mathbb{1} + (\boldsymbol{s}\cdot\hat{\mathbf{a}})(\hat{\mathbf{b}}\cdot\hat{\mathbf{a}})(\hat{\mathbf{b}}\cdot\boldsymbol{\sigma})], \tag{1.5d}$$

with respective eigenvalues

$$\lambda^{\epsilon}_{\Phi_A(\rho)} = \frac{1+\epsilon\|(\boldsymbol{s}\cdot\hat{\mathbf{a}})\hat{\mathbf{a}}\|}{2} = \frac{1+\epsilon|(\boldsymbol{s}\cdot\hat{\mathbf{a}})|}{2} \qquad \epsilon = \pm 1, \tag{1.6a}$$

$$\lambda^{\epsilon}_{\Phi_B(\rho)} = \frac{1+\epsilon\|(\boldsymbol{s}\cdot\hat{\mathbf{a}})(\hat{\mathbf{a}}\cdot\hat{\mathbf{b}})\hat{\mathbf{b}}\|}{2} = \frac{1+\epsilon|(\boldsymbol{s}\cdot\hat{\mathbf{a}})(\hat{\mathbf{a}}\cdot\hat{\mathbf{b}})|}{2} = \lambda^{\epsilon}_{\Phi_{BA}(\rho)} \qquad \epsilon = \pm 1, \tag{1.6b}$$

$$\lambda^{\epsilon}_{\Phi_{AB}(\rho)} = \frac{1+\epsilon\|(\boldsymbol{s}\cdot\hat{\mathbf{a}})(\hat{\mathbf{a}}\cdot\hat{\mathbf{b}})^2\hat{\mathbf{a}}\|}{2} = \frac{1+\epsilon|(\boldsymbol{s}\cdot\hat{\mathbf{a}})(\hat{\mathbf{a}}\cdot\hat{\mathbf{b}})^2|}{2} \qquad \epsilon = \pm 1. \tag{1.6c}$$

Thus, the relative entropies are

$$S(\Phi_A(\rho)|\Phi_{AB}(\rho)) = \left[\sum_{\epsilon}\frac{1+\epsilon|\boldsymbol{s}\cdot\hat{\mathbf{a}}|}{2}\log\frac{1+\epsilon|\boldsymbol{s}\cdot\hat{\mathbf{a}}|}{2}\right] - \left[\sum_{\epsilon}\frac{1+\epsilon|\boldsymbol{s}\cdot\hat{\mathbf{a}}|}{2}\log\frac{1+\epsilon|(\boldsymbol{s}\cdot\hat{\mathbf{a}})(\hat{\mathbf{a}}\cdot\hat{\mathbf{b}})^2|}{2}\right], \tag{1.7a}$$

$$S(\Phi_B(\rho)|\Phi_{BA}(\rho)) = 0, \tag{1.7b}$$

and subsequently, the quantifier is

$$\mathcal{I}_{\{\Phi_A(\sigma),A,B\}} = \frac{1}{2}\left\{\left[\sum_{\epsilon}\frac{1+\epsilon|\boldsymbol{s}\cdot\hat{\mathbf{a}}|}{2}\log\frac{1+\epsilon|\boldsymbol{s}\cdot\hat{\mathbf{a}}|}{2}\right] - \left[\sum_{\epsilon}\frac{1+\epsilon|\boldsymbol{s}\cdot\hat{\mathbf{a}}|}{2}\log\frac{1+\epsilon|(\boldsymbol{s}\cdot\hat{\mathbf{a}})(\hat{\mathbf{a}}\cdot\hat{\mathbf{b}})^2|}{2}\right]\right\}. \tag{1.8}$$

This equation also sheds light on how a context where A is an element of reality for ρ is not enough to reduce this definition of incompatibility to a measurement incompatibility alone since there is a lingering dependence on s, which is a vector associated with the initial state preparation.