ABOUT THE JOINT MEASURABILITY OF OBSERVABLES

ALEJANDRO AGUSTÍ MARTÍNEZ-SORIA GALLO
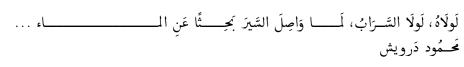
Memòria presentada per a l'obtenció dels títols de
graduat en física i graduat en matemàtiques.

Dirigit per:
Prof. Dr. Ramón Muñoz Tapia   Prof. Dr. Andreas Winter
Grup d'informació quàntica
Universitat autònoma de Barcelona

Bellaterra, 4 de setembre del 2014

CONTENTS

لَولَاهُ، لَولَا السَّـرَابُ، لَـــــــا وَاصِلَ السَّيرَ بَحِـــــثًا عَنِ الـــــــــــــــــاء ...

مَحــمُود دَرويش

Were it not for it, were it not for the mirage, [he] would not continue the journey in search for water...

Mahmoud darwish

1. INTRODUCTION

This report deals with some aspects about the joint measurability of quantum observables. Since W. Heisenberg reviewed the concepts of momentum and position, it has been known that Quantum Mechanics furnishes the impossibility of measuring them together, i.e., it is not possible to measure the one without disturbing the other. This fact holds for many other sets of observables, and it is commonly known by the name of *complementarity*.

Some of the examples taught regularly in introductory Quantum Mechanics courses also deal with these constraints in the quantum measurement, however they are rarely presented in its most generality. The mathematical model describing observables in these courses has severe limitations concerning both theoretical and experimental aspects. Even at the beginning of Quantum Mechanics there were doubts whether or not when measuring the position of some quantum particle the outcome or probability distribution was really that of the model, or otherwise some other *fuzzy* version of it. It is looking for a complete *Theory of Measurement* that many aspects discussed in this report arose for the first time. One of these new properties in the Theory of Measurement is the concept of noise, which we will see plays an important rôle in questions of measurement.

In this project we have put our focus on two outcome observables, i.e., physical measurable properties that may have as an outcome one given value or another. We have developed geometrical methods (*joint measurability graphs*) to visualize and characterize the joint measurability of any set of two outcome observables, together with constraints for noise which must be added to the system in order to make them jointly measurable.

In section 2 we will define exactly what we understand by a *general* observable. We have devoted section 3 to introduce shortly the mathematical idea of joint measurability. Section 4 constitutes the *corpus* of the project. There we discuss in full length properties for two outcome observables based on the work in [8]. We have corrected a claim appearing in [8] and we have provided refinements and possible solutions to issues caused by problems arising from the mistake.

2. POSITIVE OPERATOR VALUED MEASURES

2.1. **Projective measurements.** It has been pointed out that the best way to describe the measurement process in the quantum theory is by considering a class of positive operators. Let us illustrate some features as a way of introduction.

Let $\mathcal{H}$ be a n-dimensional Hilbert space. Suppose we have a physical observable represented by an hermitian operator A[1]. Suppose that the operator is discrete, finite and is not degenerated. Then we can find a spectral decomposition given by

$$A = \sum_{i=1}^{n} \alpha_i \, |\alpha_i\rangle\langle\alpha_i| \, ,$$
(1)

where $\{|\alpha_1\rangle, \ldots, |\alpha_n\rangle\}$ is an orthonormal basis of $\mathcal{H}$. The outcome space of a measurement of observable A is given by the set $\Omega = \{\alpha_1, \ldots, \alpha_n\}$, this is, these are the values that one can measure considering the observable A. The axioms of quantum mechanics say that if a measure appartus happens to measure α_i on a state $|\psi\rangle$, then the state immediately after the measurement is $|\alpha_i\rangle$. It is also stated that the probability of measuring α_i on a state $|\psi\rangle$ is

$$p(\alpha_i) = |\langle\alpha_i|\psi\rangle|^2 = \langle\psi|\alpha_i\rangle\langle\alpha_i|\psi\rangle = \mathrm{tr}\left\{|\alpha_i\rangle\langle\alpha_i|\,|\psi\rangle\langle\psi|\right\}.$$

If we use the formulation considering statistical mixtures then we would have a trace one positive operator $\rho \in \mathcal{L}(\mathcal{H})$ as the state of the system and the probability of measuring α_i would then be

$$p(\alpha_i) = \mathrm{tr}\left\{|\alpha_i\rangle\langle\alpha_i|\,\rho\right\}.$$

Analogously, the probability of measuring α_i or α_j, i.e., the probability that the measurement lies within the set $\{\alpha_i, \alpha_j\}$, provided that $i \neq j$ is

$$p(\{\alpha_i, \alpha_j\}) = \mathrm{tr}\left\{\left(|\alpha_i\rangle\langle\alpha_i| + |\alpha_j\rangle\langle\alpha_j|\right)\rho\right\} = p(\alpha_i) + p(\alpha_j).$$

By considering these last equations, one can see that in fact we have defined a probability measure on 2^{Ω}, this being the set of the parts of Ω. Therefore, for every element X of 2^{Ω} we can assign the probability $p(X)$ as before.

―――――――――

[1]For simplicity's sake we do not differentiate between an observable or physical property and its assigned operator.

Altogether we actually see that as far as the measurement is concerned, given a Hilbert space $\mathcal{H}$, all the information of the observable's measurement is encoded in the outcome space Ω and the set of operators $\{|\alpha_i\rangle\langle\alpha_i| \mid i \in \{1, \ldots, n\}\}$. So the observable is an association of a subset of Ω (or an element of 2^Ω) with some operators acting on $\mathcal{H}$, which in this case it would translate into

$$\{\alpha_{i_1}, \ldots, \alpha_{i_r}\} \longmapsto |\alpha_{i_1}\rangle\langle\alpha_{i_1}| + \cdots + |\alpha_{i_r}\rangle\langle\alpha_{i_r}|.$$

It is also worth noticing that for the measurement of probabilities only a certain kind of operators are needed. Indeed operators needn't be projectors but just positive, so that the expected value is always positive and it makes sense to extract from every operator a probability value.

2.2. **POVM.** We will define what we understand by an *observable* following the footsteps and hints of section 2.1. So let $\mathcal{H}$ be a Hilbert space and let $\rho \in \mathcal{L}(\mathcal{H})$ be a state, i.e., a linear operator acting on $\mathcal{H}$ having trace one. From now on, Ω will be an outcome space associated with an observable and $\mathcal{F}$ the possible combinations of elements of this Ω we can calculate probabilities of, i.e., $\mathcal{F}$ is a σ-Algebra attached to Ω.

Definition 2.1. *Let $(\Omega, \mathcal{F})$ be as above. A map $\mathsf{A} : \mathcal{F} \to \mathcal{L}(\mathcal{H})$ is an operator valued measure if and only if the following statements are true:*

(1) $\mathsf{A}(\emptyset) = 0$.

(2) *Let $\{X_i \mid i \in \mathbb{N}\}$ be a countable subset of $\mathcal{F}$. If $X_i \cap X_j = \emptyset$ for $i \neq j$ and $i, j \in \mathbb{N}$ then*

$$\mathrm{tr}\left\{\mathsf{A}\left(\bigcup_{i \in \mathbb{N}} X_i\right)\rho\right\} = \sum_{i \in \mathbb{N}} \mathrm{tr}\left\{\mathsf{A}(X_i)\rho\right\}.$$

Note that for the case that the dimension of the Hilbert space is 1, the above definition is just the definition of a measure.

Definition 2.2. *Inside the framework of definition 2.1:*

(1) *an operator valued measure $\mathsf{A} : \mathcal{F} \to \mathcal{L}(\mathcal{H})$ is called an* observable *if and only if $\mathsf{A}(X)$ is positive semi-definite for all $X \in \mathcal{F}$ and $\mathsf{A}(\Omega) = 1$. We refer to them also as positive operator valued measure or in short POVM.*

(2) *an observable is called* sharp *if and only if $\mathsf{A}(X)$ is a projector for every $X \in \mathcal{F}$, i.e., $\mathsf{A}(X)^2 = \mathsf{A}(X)$.*

These last definitions are quite general, maybe even they are too general for the purposes of this report, however in proposition 4.10 we make use of this general framework to prove a proposition. Let show this definition in the simple case of a discrete and finite outcome space Ω. In this case $\Omega = \{\alpha_1, \ldots, \alpha_n\}$ and since the nature of α_k is not important for the purpose of measuring, we may as well call α_k simply k, so that $\Omega = \{1, \ldots, n\}$. In this case we have for every element of Ω a positive operator $A_i \stackrel{\text{def}}{=} \mathsf{A}(i)$. So we have a set of positive operators $\{A_i, i \in \{1, \ldots, n\}\}$ which fulfill the condition

$$\sum_i A_i = 1$$

for being A and observable. In the case of a sharp observable, every A_i is of the form $|i\rangle\langle i|$ for some $|i\rangle \in \mathcal{H}$.

Sharp observables of course are meant to be the self-adjoint operators representing observables that are commonly introduced in Quantum Mechanics courses. Returning to the example of last section with the operator A in equation 1, we made the case that for every possible outcome α_j we had the projector $|\alpha_j\rangle\langle\alpha_j|$ that would give us the propability of measuring α_j in the sate $|\phi\rangle$ by $\mathrm{tr}\{|\alpha_j\rangle\langle\alpha_j| \, |\phi\rangle\langle\phi|\}$. Theoretically this class of observables has a very important property, which has to do with the name *sharp*. Let us suppose we are measuring A on a system which is in the state $|\phi\rangle = |\alpha_j\rangle$, since A is sharp, it is composed from projectors and we suppose that all projectors are orthogonal, then we have the following condition on the probability of $|\phi\rangle$

$$(2) \qquad \forall i \quad \mathrm{tr}\{A(\alpha_i) \, |\phi\rangle\langle\phi|\} = \delta_{ij}.$$

In this case, for this particular state $|\phi\rangle$ the distribution is as sharp as a distribution can get. In general sharp observables can get arbitrarily sharp, like the position observable, where we can perform experiments to lessen the uncertainty σ_x.

However imagine we do not know the start state $|\phi\rangle$ or this state is in a statistical mixture ρ such that $\rho^2 \neq \rho$, i.e., such that $\rho \neq |\psi\rangle\langle\psi|$ for any $|\psi\rangle$. In this case it is not anymore possible to find a ρ such that $\rho^2 \neq \rho$ having the property of equation 2.

In this way POVM's act as real measurements [5] since they do not have the *sharpness* property. Of course if we consider Quantum mechanics to be a complete

theory of measurement, taking into consideration the dimension d of the Hilber Space $\mathcal{H}$, if we were to content ourselves with *sharp* observables we would not be able to consider measurements with more than d outcomes, which in the framework of experimentation would be a dramatic hindrance.

POVM's give therefore the idea that we should be able to consider any probabilistic-valid distribution we want. They are indeed objects that cater an infinity of probability distributions by definition. This means, more concretely, if $(\mathsf{A}, \Omega_A, \mathcal{F}_A)$ is a POVM, that for every ρ we get a probability distribution over $\mathcal{F}_A$. Indeed, we consider the function that to every $U \in \mathcal{F}_A$ assigns the number $\mathrm{tr}\{A(U)\rho\}$. This function is by definition of A a probability distribution over $\mathcal{F}_A$.

2.3. **Neumark's theorem.** POVM arise naturally from bipartite systems. In essence, the measurement process is a coupling between the system to be measured A and the system measuring B [5]. The reading process takes place in B, and it is an irreversible process. This means conceptually that reading the information in B after the coupling may cease to be a *sharp* reading, i.e., when one looks at the probability distribution of the readings in B, the structure does not comply with the one of a sharp observable.

However, there is a way around, which is known as *Neumark's theorem* at it is stated below [1].

Theorem 2.1. *Let A be a POVM over a given Hilbert space $\mathcal{H}$. Then there exists a sharp observable A^+ on an extended Hilbert space $\mathcal{H}^+$ such that if $P : \mathcal{H}^+ \to \mathcal{H}$ is the projection from the extended Hilbert space to the former then for every $|\psi\rangle \in \mathcal{H}$*

$$A |\psi\rangle = P A^+ |\psi\rangle .$$

In simple words what the theorem states is that every POVM can actually be thought as a sharp observable on a bigger Hilbert space. Therefore the nature of POVM's gets clarified by it, since any POVM is either sharp or is a partial trace of a *bigger* sharp observable.

3. Join-mesurability of POVM's

3.1. Definition and properties.

Definition 3.1 (joint measurability). *Let* $\{A_i \mid i \in \mathcal{N} = \{1,\dots,n\}\}$ *be a set of POVM's with outcome spaces* Ω_{A_i} *and* σ*-algebras* $\mathcal{F}_{A_i}$. *We say that they are jointly measurable if there exists a POVM R with outcome space* $\Omega_{A_1} \times \cdots \times \Omega_{A_n}$ *and* σ*-algebra* $\mathcal{F}_{A_1} \otimes \cdots \otimes \mathcal{F}_{A_n}$ *such that every* A_i *is a marginal of R, i.e., for every* $i \in \mathcal{N}$ *and every* $U \in \mathcal{F}_{A_i}$

$$A_i(U) = Z(\Omega_{A_1} \times \cdots \Omega_{A_{i-1}} \times U \times \Omega_{A_{i+1}} \times \cdots \times \Omega_{A_n})$$

Let us illustrate this somewhat general definition with an example of sharp observables. Suppose therefore that

$$\hat{A} = \sum_i \alpha_i \lvert\alpha_i\rangle\langle\alpha_i\rvert, \qquad \hat{B} = \sum_i \beta_j \lvert\beta_i\rangle\langle\beta_i\rvert$$

are two self-adjoint operators. To build the POVM's out of the two self-adjoint operators we consider the outcome spaces $\Omega_A = \{\alpha_i\}$ and $\Omega_B = \{\beta_i\}$ and secondly we consider the map

$$A : \{\alpha_{i_1}, \dots, \alpha_{i_n}\} \mapsto \lvert\alpha_{i_1}\rangle\langle\alpha_{i_1}\rvert + \cdots + \lvert\alpha_{i_n}\rangle\langle\alpha_{i_n}\rvert.$$

Let us call this map A (resp. B). As it is known from Quantum Mechanics, $\hat{A}$ and $\hat{B}$ are called jointly measurable if and only if they commute since only then we can find a common eigenbasis for both operators. Therefore we could find vectors $\lvert\alpha_i\beta_j\rangle$ such that $\hat{A}\lvert\alpha_i\beta_j\rangle = \alpha_i\lvert\alpha_i\beta_j\rangle$ and $\hat{B}\lvert\alpha_i\beta_j\rangle = \beta_j\lvert\alpha_i\beta_j\rangle$, which it is interpreted as that the state $\lvert\alpha_i\beta_j\rangle$ has both the property α_i and β_j *at the same time.*In this sense, only considering commutativity it is meaningful to consider both properties at the same time.

Consider now a pure state $\lvert\psi\rangle$, the probability of measuring α_i is of course given by $\lvert\langle\alpha_i\vert\psi\rangle\rvert^2 = \mathrm{tr}\{\lvert\alpha_i\rangle\langle\alpha_i\rvert\,\lvert\psi\rangle\langle\psi\rvert\}$ and after having measured α_i, the state ψ collapses irreversibly to α_i. If we perform another measurement B and we measure β_j, the probability for it is $\mathrm{tr}\{\lvert\alpha_i\rangle\langle\alpha_i\rvert\,\lvert\beta_j\rangle\langle\beta_j\rvert\}$.

If we consider therefore the POVM $Z = A \cdot B$ having as outcome space $\Omega_A \times \Omega_B$ and as σ-algebra $\mathcal{F}_A \otimes \mathcal{F}_B$, it fulfills the following relations:

$$A(\alpha_i) = Z(\alpha_i, \Omega_B) = \sum_{\beta \in \Omega_B} Z(\alpha_i, \beta), \quad B(\beta_j) = Z(\Omega_A, \beta) = \sum_{\alpha \in \Omega_A} Z(\alpha, \beta_j)$$

so we could think that we have found the joint measure for A and B. But, notice that we have not used anywhere the fact that $\hat{A}$ and $\hat{B}$ are commute. So we could be falsely led to think that everything is jointly measurable! However, that is not the case. In general nothing ensures the semidefinite positivity of

$$Z(\alpha, \beta) = |\alpha\rangle\langle\alpha| \, |\beta\rangle\langle\beta|$$

even though both $|\alpha\rangle\langle\alpha|$ and $|\beta\rangle\langle\beta|$ are positive. It is a theorem of the functional analysis [6] that the product of positive operators is ensured to be positive only if they commute. Therefore Z would be a joint measure if all projectors commute which implies the commutation of $\hat{A}$ and $\hat{B}$.

3.2. **Adding noise.** Given two non-jointly measurable POVM's A and B, it is possible to make them joint measurable by *adding* noise into them? What do we understand by noise? Let us illustrate what is meant by it through an example. Suppose the outcome space of A is $\{a_1, \ldots, a_n\}$ and suppose we perform an experiment to measure A several times, say M times where M should be big enough. Suppose the measurement gets each time states ρ which are assumed to be equally prepared. Then considering the relative frequencies of every $a_i \in \Omega_A$, i.e. the times we have measured a_i divided by M, if M is big enough we could identify this relative frequency $f(a_i)$ with the probability, i.e.,

$$f(a_i) \approx \text{tr}\{\rho A(a_i)\}$$

and we could have a graph such as in figure 1 in blue.

In practice however we never get such a distribution, or we are not sure to have the real distribution, so one says that there is noise in the system. The measuring process is a macroscopic process between some quantum system and a bigger one, where a thermodynamic irreversible process takes action (see [4] or [5]). In this process many effects may take place leading to mistakes in the reading of some sensor. These factors are noise sources, and we can mathematically model this influence by taking a convex combination of POVM's, in the case of A for example we could take another POVM E having too as outcome space Ω_A and representing
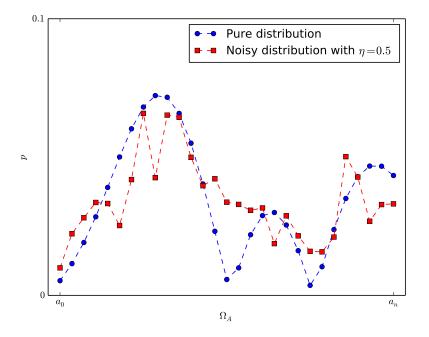
FIGURE 1. Graphical representation of the distribution of a given observable A (in blue) compared with some noisy distribution of the same A(in red). The lines have been added to give also a feeling of how a noisy distribution in the case of a continuous observable would look like.

the noise. A convex combination is provided taking a parameter $\eta \in [0, 1]$ and considering the new POVM

$$\mathsf{A}_\eta = (1 - \eta)\mathsf{A} + \eta E.$$

In figure 1 we can see a representation of how some results would look like if we were to note down the results of the measurements on some random state ρ and if we could compare the noisy distribution with the pure one. We see that as an overall effect what this tells us is that we actually make gradually less sharp the information until all the information that is left is the unusable information from E, which again might be a random noise $POVM$. Therefore it is not strange to assume that considering some η we could make A_η and B_η joint measurable, since we lose more and more information as η increases. Let us stress that such convex combinations are not a full description of the noise, but a model. For example note that we can never attain from a positive operator $\mathsf{A}(a_i)$ to make it zero using such a combination, since $\mathsf{A}_\eta(a_i) > 0$ if $\mathsf{A}(a_i) > 0$.

4. Joint measurablility of two-outcome observables

Nomenclature: Throughout this section we will call two outcome observables *effects* interchangeably, although in the literature we find different uses for the word effect. Also, given a positive operator P, we may refer to it as an observable, meaning the two outcome observable $\mathsf{P} = \{P, \mathbb{1} - P\}$, it is assumed that $P \leq \mathbb{1}$. Both P and $\mathbb{1} - P$ are called *POVM elements*, since they conform together the two outcome POVM or observable P. We will normally note the outcome space of two outcome observables as $\Omega = \{+, -\}$, so that for example $\mathsf{P}(+) = P$ and $\mathsf{P}(-) = \mathbb{1} - P$, or otherwise. The association of $+$ to P or $\mathbb{1} - P$ is but a matter of preference and language, the physical significance is encoded in P.

4.1. **Introduction.** Some conditions for the joint measurability of effects have already been found in the literature in the form of inequalities of a certain set of operators (see [8]). In what follows we will attempt to present a generalization and correction of some aspects of these results.

The description of the process needs of an explanation of the framework. In the case of $n = 2$, one can find quite straightforward methods to check and find conditions for their joint measurability. Let $\{P, \mathbb{1} - P\}, \{Q, \mathbb{1} - Q\}$ be two observables having both the outcome space $\Omega_{P,Q} = \{+, -\}$. If they are jointly measurable, a set of operators exists $R(i, j)$, $i, j \in \{+, -\}$, for which the coarse-graining condition is satisfied, i.e.

$$P = R(+, +) + R(+, -), \qquad Q = R(+, +) + R(-, +)$$

$$\mathbb{1} - P = R(-, +) + R(-, -), \qquad \mathbb{1} - Q = R(-, +) + R(-, -)$$

The condition proven in [8] states that P and Q are jointly measurable if and only if there exists a positive operator S such that the following requirements are satisfied,

$$(3) \qquad\qquad P + Q - \mathbb{1} \leq S, \qquad S \leq P, Q.$$

It is a matter of calculation to check the necessity of the condition, simply take *an operator of the R-representation which is common to both P and Q*, and this will do the trick (i.e., take $R(+, +)$). The sufficiency condition hides some interesting steps; given such an S, the conditions $S \leq P, Q$ allow us to define $R(+, -) \stackrel{\text{def}}{=} P - S$ and $R(-, +) \stackrel{\text{def}}{=} Q - S$, which by hypothesis are positive. The condition on the left allows us to define both $R(-, -) \stackrel{\text{def}}{=} \mathbb{1} - P - Q + S \geq 0$ (which will be positive) and

to establish the validity of the partition of unity condition of R, i.e., $1 = \sum_{i,j} R(i,j)$ where $i, j \in \{-, +\}$. Therefore by calling $S = R(+, +)$ we encounter the desired joint-measure.

In [8], the general conditions for n two outcome observables are written down in terms of the joint measure R directly. However it is of our opinion that to research some properties of n observables, this method is quite opaque in the reading. We have developed a general method using a generalization of the operator S rather than R directly. However, as dimension increases the number of S's increases as well and the huge number of inequalities ceases to have the helpful factor it had in the latter simpler case. It is therefore difficult to see properties or possible solutions to problems arising from the joint measurability through the pure inequality framework. Therefore, we have thought of these inequalities as conforming a directed graph with several floors or levels, and we have derived some general properties of these graphs relevant to physical problems. The direction of the graph is intended to give us some hint about the underlying inequalities.

We will discuss the structure of the graph defined for the case of $n = 2$ later on, when we prove the equivalence theorem in the following sections (theorem 4.1). There we prove that the considerations above for two observables and for n observables in [8] are equivalent to considering the graph schema:
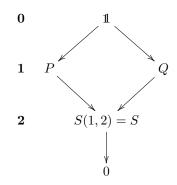

FIGURE 2. Directed graph for two effects.

The fact that for higher dimension (i.e., more than two effects) we do not only have an operator S to determine, makes complicated to *parametrize* these operators. We have to be able to write them in an operational and ordered way so that we can identify exactly each operator and we can eventually write algorithms to solve the

problems. We use a subset approach to the problem, i.e., every operator and the inequalities are defined in terms of subset properties as it will be seen in definition 4.1, where we define formally the properties of the graphs and we call them *joint measurability* graphs for the sake of readability of the report.

4.2. **Definition of the graph.** From now on let us define $\mathcal{N} = \{1, \ldots, n\}$.

Definition 4.1 (Joint measurability graph)**.** *Let $G = (V, A)$ be a directed graph. The graph G is called a joint measurability graph of dimension n iff the following conditions are met:*

(1) *V is a set of operators acting on a Hilbert space.*

(2) *If $(S_1, S_2) \in A$ then $S_1 \geq S_2$.*

(3) *There exists a bijection*

$$S : 2^{\{1, \ldots, n\}} \to V$$

(4) *The bijection defines the elements in A,*

$$A = \{(S(\Omega_1), S(\Omega_2)) \mid \Omega_1 \subset \Omega_2 \text{ and } |\Omega_2 \backslash \Omega_1| = 1\}$$

(5) *For every Ω, $\Omega \subseteq \mathcal{N}$ the following defining inequality is satisfied:*

$$(4) \qquad R(\Omega) \stackrel{\text{def}}{=} \sum_{A \subseteq \mathcal{N} \backslash \Omega} (-1)^{|A|} S(\Omega \cup A) \geq 0$$

Please note that point 2 is a consequence of point 5 although this can not be seen easily directly. However it is desirable to include point 2 in the definition to make it more transparent and readable. Equation 4 may seem a little bit difficult to read at first. Despite this fact, it will become clear as we show the necessity of such a graph in theorem 4.1. For now let us write an example with dimension 3. In this case we have $\mathcal{N} = \{1, 2, 3\}$ and $2^{\mathcal{N}} = \{\emptyset, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}\}$. We write for simplicity $S(1, 3)$ instead of $S(\{1, 3\})$. Let us write equation 4 for $\Omega = \{1, 2, 3\}$, since in this case $\Omega = \mathcal{N}$ the only possible $A \subset \mathcal{N} \backslash \Omega$ will be $A = \emptyset$, so in this case it reads:

$$R(\mathcal{N}) \stackrel{\text{def}}{=} (-1)^0 S(\mathcal{N} \cup \emptyset) = S(\mathcal{N}) \geq 0.$$

For $\Omega = \{1, 3\}$ for example, $\mathcal{N} \backslash \Omega = \{2\}$, so A can be $\emptyset$ and $\{1\}$, let us write it:

$$R(1, 3) = (-1)^0 S(\{1, 3\} \cup \emptyset) + (-1)^1 S(\{1, 3\} \cup \{2\}) = S(1, 3) - S(\mathcal{N}) \geq 0$$

and of course the same is valid for $\{1, 2\}$ and $\{2, 3\}$. From here we get therefore that $S(A_2) \geq S(1, 2, 3) \geq 0$ where of course A_2 is a subset of $\mathcal{N}$ of cardinality two. Now suppose $\Omega = \{1\}$, then A can be $\{2\}$, $\{3\}$, $\{2, 3\}$ or $\emptyset$, equation 4 turns therefore into

$$S(1) - S(1, 3) - S(1, 2) + S(\mathcal{N}) \geq 0$$

where we can write the same (but slightly different) for $S(2)$ and $S(3)$. Finally for $\Omega = \emptyset$ we can write

$$S(\emptyset) - \sum_{i=1}^{3} S(i) + S(1,2) + S(1,3) + S(2,3) - S(1,2,3) \geq 0.$$

All these inequalities can be represented in form of a graph. In figure 3 we can see the representation of such a directed graph.
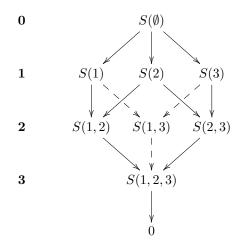

FIGURE 3. Joint measurability graph for $n = 3$.

We can find a thumb-rule to find the inequalities without having to look at the subsets (compare also with section 4.1). Just take any element $S(\Omega)$ and look at the storey it is in, i.e., **0, 1, 2** or **3**. Then substract everything connected to $S(\Omega)$ lying in the storey immediately beneath, then add everything connected to the elements substracted of the next storey to them, etc. . . Let us illustrate through an example; take for instance $S(2)$ which is in the storey **1** (because $\{2\}$ has one element), connected to it beneath are $S(1,2)$ and $S(2,3)$, so we have to substract them, i.e., $S(2) - S(1,2) - S(2,3)$. We are not yet quiet there, we have to add the element connected to $S(1,2)$ and $S(2,3)$, which is only $S(1,2,3) = S(\mathcal{N})$. Furthermore being a joint measurability means that the result must be greater or equal zero, so

$$S(2) - S(1,2) - S(2,3) + S(1,2,3) \geq 0.$$

For the following considerations the following lemma will prove to be useful.

Lemma 4.1. *Let $\mathcal{N} = \{1, \ldots, n\}$ and $\Omega \subseteq \mathcal{N}$. Let $S, R : 2^{\mathcal{N}} \to \mathcal{R}$ be two functions taking values in any $\mathbb{C}$ ring $\mathcal{R}$. The two functions fulfill the relation*

$$(5) \qquad R(\Omega) = \sum_{A \subseteq \mathcal{N} \setminus \Omega} (-1)^{|A|} S(\Omega \cup A)$$

if and only if they also fulfill

$$(6) \qquad S(\Omega) = \sum_{A \subseteq \mathcal{N} \setminus \Omega} R(\Omega \cup A)$$

which can be understood as a coarse-grain property.

Proof. We begin by showing that equation 5 implies 6. Let us denote the right-hand side of equation 6 by Γ. By definition, every term appearing in the summation of Γ is

$$R(\Omega \cup A) = \sum_{A' \subseteq \mathcal{N} \setminus (\Omega \cup A)} (-1)^{|A'|} S(\Omega \cup A \cup A')$$

and with it we can write for Γ

$$
\begin{aligned}
\Gamma &= \sum_{A \subseteq \mathcal{N} \setminus \Omega} \sum_{A' \subseteq \mathcal{N} \setminus (\Omega \cup A)} (-1)^{|A'|} S(\Omega \cup A \cup A') \\
&= \sum_{A' \subseteq \mathcal{N} \setminus \Omega} (-1)^{|A'|} S(\Omega \cup A') + \sum_{\substack{A \subseteq \mathcal{N} \setminus \Omega \\ |A| \neq 0}} \sum_{A' \subseteq \mathcal{N} \setminus (\Omega \cup A)} (-1)^{|A'|} S(\Omega \cup A \cup A') \\
&= R(\Omega) + \sum_{\substack{A \subseteq \mathcal{N} \setminus \Omega \\ |A| \neq 0}} \sum_{A' \subseteq \mathcal{N} \setminus (\Omega \cup A)} (-1)^{|A'|} S(\Omega \cup A \cup A')
\end{aligned}
$$

where we have separated the summation by taking $A = \emptyset$ on one and $A \neq \emptyset$ on the other, we have used the definition of $R(\Omega)$. The second last summation depends therefore on A and A', where A' has in its turn also a dependence on A. We will change of variable for the summation by taking

$$Z = A \cup A'$$

and finding a summation over Z. Let us note however what are the ranges of $|A|$, $|A'|$ and $|Z|$ all together:

$$
\left.
\begin{aligned}
|A| &\in & \{1, \ldots, n - |\Omega|\} \\
|A'| &\in & \{0, \ldots, n - |\Omega| - |A|\}
\end{aligned}
\right\} \Rightarrow |Z| \in \{1, \ldots, n - |\Omega|\}
$$

From the latter consideration it follows that Z can be any subset of $\mathcal{N} \setminus \Omega$ with cardinality greater or equal than 1. Let us consider a given Z, there are several

forms of obtaining it from sets A and A'. For example if $|A| = 1$, A' is automatically determined by $Z \backslash A$ and there are $\binom{|Z|}{1}$ ways of doing it, which correspond of course to the possible subsets of Z of length 1. Writing the possibilities in dependence of the cardinality of A there are therefore

$$\sum_{i=1}^{|Z|} \binom{|Z|}{i} = 2^{|Z|} - 1$$

possible ways of producing $Z = A \cup A'$, what means that $S(\Omega \cup Z)$ appears $2^{|Z|} - 1$ times in $\Gamma - R(\Omega)$. In order to sum up all these terms we have to consider the change of sign according to the cardinality of A', which is $|Z| - |A|$ or in the sum above $|Z| - i$. We can therefore write

$$\sum_{i=1}^{|Z|} (-1)^{|Z|-i} \binom{|Z|}{i} S(\Omega \cup Z) = [(1-1)^{|Z|} - (-1)^{|Z|}] S(\Omega \cup Z) = -(-1)^{|Z|} S(\Omega \cup Z)$$

This means however that we can rewrite Γ simply in terms of Z in the following way

$$\Gamma = R(\Omega) + \sum_{\substack{Z \subseteq \mathcal{N} \backslash \Omega \\ |Z| \neq 0}} (-1)(-1)^{|Z|} S(\Omega \cup Z)$$

$$= R(\Omega) - \sum_{\substack{Z \subseteq \mathcal{N} \backslash \Omega \\ |Z| \neq 0}} (-1)^{|Z|} S(\Omega \cup Z) = S(\Omega)$$

which is the exactly $S(\Omega)$ by definition on equation 4.

Implication of equation 5 from equation 6 is shown similarly. We denote as before the right-hand side of equation 5 as Γ and express every term of the sum in function of the assumption:

$$\Gamma = \sum_{A \subseteq \mathcal{N} \backslash \Omega} \sum_{A' \subseteq \mathcal{N} \backslash (\Omega \cup A)} (-1)^A R(\Omega \cup A \cup A')$$

$$= R(\Omega) + \sum_{\substack{A \subseteq \mathcal{N} \backslash \Omega \\ |A| + |A'| \neq 0}} \sum_{A' \subseteq \mathcal{N} \backslash (\Omega \cup A)} (-1)^A R(\Omega \cup A \cup A')$$

where of course we have put out of the summation the pair (A, A') where both of them where the emptyset. Now the only constraint on the sets $Z = A \cup A'$, A and A' is that A and A' can not be the empty set at the same time. Following the reasoning of before and for given $Z = A \cup A'$ we find that the sum of the elements

$R(\Omega \cup Z)$ is

$$(-1)^0 \binom{|Z|}{0} + (-1)^1 \binom{|Z|}{1} + \ldots + (-1)^{|Z|} \binom{|Z|}{|Z|} = (1-1)^{|Z|} = 0.$$

As before, writing the equation changing variables with Z we find that $\Gamma = R(\Omega)$ as desired.

$\square$

4.3. **Graph equivalence.** We set out to prove the equivalence of *joint measurability* graphs and joint measurable effects. For that we prove the following main theorem:

Theorem 4.1. *Let* $\{P_1, \ldots, P_n\}$ *be a set of effects. These effects are jointly-measurable if and only if there exists an n-dimensional joint measurability graph* (V, A, S) *such that* $S(i) = P_i$ *for every* $i \in \{1, \ldots, n\}$ *and* $S(\emptyset) = \mathbb{1}$.

Proof. Let us begin by showing the **necessity**. Let us suppose the effects take values on $\{+, -\}$, so that P_i is the operator related to $+$. If they are joint measurable then there exists a POVM $R : \{+, -\}^n \to \mathcal{L}(\mathcal{H})$ fulfilling

$$P_i = \sum_{\substack{t_i = + \\ t_k \in \{+, -\}}} R(t_1, \ldots, t_i, \ldots, t_n).$$

For the sake of readability let us denote, $R(t_1, \ldots, t_n)$, where some of the t_k are equal to $+$, by this subset of indexes, i.e., we do not write the indexes for which $t_k = -$, e.g. :

$$R(+, -, +, -, \ldots, \overset{k}{+}, -, \ldots, +) =: R(\{1, 3, k, n\}) = R(1, 3, k, n).$$

Let $\mathcal{N} = \{1, \ldots, n\}$ and S be a function over $2^{\mathcal{N}}$ defined as

$$S(\Omega) = \sum_{A \subset \mathcal{N} \backslash \Omega} R(\Omega \cup A), \quad \forall \Omega \subseteq \mathcal{N}.$$

Since $\mathcal{N}$ is finite and $R(A)$ is well-defined for every $A \subset \mathcal{N}$, S is well-defined and finite. Let us remark that this is in fact a mere general coarse-graining property since as a consequence of the definition $S(i) = P_i{}^2$ and $S(\emptyset) = \mathbb{1}$. Consider $V = \{S(\Omega) \mid \Omega \subseteq \mathcal{N}\}$ and $A = \{(S(\Omega_1), S(\Omega_2)) \mid \Omega_1 \subseteq \Omega_2, |\Omega_2 \backslash \Omega_1| = 1\}$, we claim that with it (A, V, S) is a joint measurability graph. Since we have defined A and V through S there are only two conditions left to be shown, i.e.:

- $(S(\Omega_1), S(\Omega_2)) \in V \Rightarrow S(\Omega_1) \geq S(\Omega_2)$: This is the case since every term appearing in definition of $S(\Omega_2)$ appears also on the definition of $S(\Omega_1)$ due to $\Omega_1 \subseteq \Omega_2$.
- That

$$R(\Omega) = \sum_{A \subseteq \mathcal{N} \backslash \Omega} (-1)^{|A|} S(\Omega \cup A)$$

  is a consequence of lemma 4.1. Moreover, the expression is positive semi-definite for all $\Omega \subseteq \mathcal{N}$ by hypothesis.
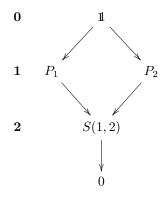
---

[2]For simplicity we write $S(i_1, \ldots, i_t)$ instead of $S(\{i_1, \ldots, i_t\})$

Not we proceed to show the **sufficiency**. Suppose therefore that there exists a joint measurability graph (V, A, S) meeting the requirements of the hypothesis. The joint measurablility of the effects follows directly from the definition of the graph. Let us define $R : \{+, -\}^n \to \mathcal{L}(\mathcal{H})$ as before, namely from $R(\Omega)$ of the graph we define $R(+, -, -, \ldots)$ with $+$ in the positions determined by the elements of Ω. They are positive semi-definite and from lemma 4.1 we have

$$P_i = \sum_{A \subset \mathcal{N} \setminus \{i\}} R(\{i\} \cup A) = \sum_{\substack{t_i = + \\ t_k \in \{+, -\}}} R(i_1, \ldots, t_{i-1}, +, t_{i+1}, \ldots, t_n)$$

which gives us the coarse graining condition. The σ-additivity of R is automatic since $\{+, -\}^n$ is finite and R is defined through single elements of the set.

$\square$

We apply the theorem 4.1 to the case $n = 2$ as an example. Let us therefore consider two jointly measurable effects $\{P_1, \mathbb{1} - P_1\}$ and $\{P_2, \mathbb{1} - P_2\}$. In this case we have the function $S : 2^{\{1,2\}} \to V$ where $S(\emptyset) = \mathbb{1}$ and $S(i) = P_i$. We have therefore a $S(1, 2)$ such that the following graph is satisfied:

As expressed in the theorem, if P_i is associated with the value $+$ and $\mathbb{1} - P_i$ with $-$, then for instance $R(1, 2)$ means $R(++)$. In this case $0 \leq R(++) = S(1, 2)$ and

$$R(1) = P_1 - S(1, 2) = P_1 - R(++) = R(+-) \geq 0, \qquad P_2 - R(++) = R(-+) \geq 0$$

and finally

$$R(\emptyset) = \mathbb{1} - P_1 - P_2 + R(++) = R(--) \geq 0.$$

These $R(i, j) \geq 0$ form indeed a joint-measure of P_1 and P_2, which is exactly the result explained in section 4.1. In this way, theorem 4.1 can be thought of as a generalization of Proposition 1 in [8].

4.4. **Graphs as semidefinite programs.** As shown in [8] the joint measurablil-ity problem of observables can be cast in the form of a semidefinite program. We show this same process making use of the graph structure. This is helpful when there is the need of performing a simulation of the joint measurablility of effects or to specify graphically and clearly some properties of these so created semidefinite programs.

The primal problem of a *semidefinite program* (SDP) (see [7]) consists in mini-mizing a linear function of a vector $\vec{x} \in \mathbb{R}^m$ subjected to a linear matrix inequality constraint, i.e.

$$(7) \qquad \min_{\vec{x}} \left\{ c^T x \,\middle|\, F_0 + \sum_{i=1}^{m} x_i F_i \geq 0 \right\}$$

where F_i are n-dimensional symmetric matrices. The data of the problem is therefore $c \in \mathbb{R}^m$ and $\{F_i \mid i \in \{0, \dots, m\}\} \subset M_n(\mathbb{R})$. This kind of optimization problems is efficiently solvable using methods like *interior-point methods* which are generally used also for linear programs.

Let $\{P_1, \dots, P_n\}$ be n effects and $G = (V, A, S)$ a potential joint measurability graph that would be associated to the set of effects if these were jointly measurable. Let $\mathcal{N} = \{1, \dots, n\}$ and $\{Q_i \mid i \in \{1, \dots, d\}\}$ be a hermitian basis for the space where the effects are in. Then for every $S(A)$ where $A \subset \mathcal{N}$ and $|A| \geq 2$ we write

$$S(A) = \sum_{i=1}^{d} x_i^A Q_i.$$

In this way we define variables x_i^A for every A such that $|A| \geq 2$. To refer generally to all of them it is important, specially when programming an algorithm, to define any bijection like

$$\varphi : \{1, \dots, 2^n - n - 1\} \longrightarrow \{A \subset \mathcal{N} \mid |A| \geq 2\}$$

which exists since the cardinality of the set on the right is exactly $2^n - n - 1$. Let us define $\vec{x} = (\vec{x}_1, \dots, \vec{x}_{2^n - n - 1})$ where every $\vec{x}_i$ is associated by the bijection φ to an element $\varphi(i) = A$ such that

$$\vec{x}_i = (x_1^{\varphi(i)}, \dots, x_m^{\varphi(i)}).$$

For every element $S(\Omega)$ with $|\Omega| \geq 2$ we write the graph condition as

$$\sum_{A \subseteq \mathcal{N} \setminus \Omega} (-1)^{|A|} S(\Omega \cup A) = \sum_{A \subseteq \mathcal{N} \setminus \Omega} \sum_{i=1}^{d} x_i^{\Omega \cup A} (-1)^{|A|} Q_i \geq 0$$

If $\Omega = \{k\}$ then we may write

$$P_k + \sum_{A \subseteq \mathcal{N} \setminus \{k\}} \sum_{i=1}^{d} x_i^{\{k\} \cup A} (-1)^{|A|} Q_i \geq 0$$

and finally if $\Omega = \emptyset$ we change slightly the property of the graph (i.e. $S(\emptyset) = \mathbb{1}$) and set instead $S(\emptyset) = \lambda \mathbb{1}$, where $\lambda \in \mathbb{R}^+$, thus

$$\lambda \mathbb{1} - \sum_{i \in \mathcal{N}} P_i + \sum_{\substack{A \subseteq \mathcal{N} \\ |A| \geq 2}} \sum_{i=1}^{d} x_i^{\Omega \cup A} (-1)^{|A|} Q_i \geq 0.$$
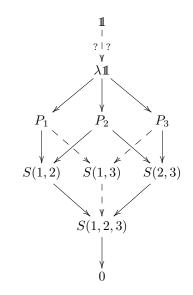

FIGURE 4. joint measurability graph representation for minimization of SDP-parameter λ.

These inequalities have the following form

$$R(\Omega) = F_0(\Omega) + \sum_{i=0}^{d(2^n - n - 1)} F_i(\Omega) x_i \geq 0, \qquad \vec{x} = (x_0, x_1, \ldots, x_{d(2^n - n - 1)})$$

which is a constraint of the form of the ones appearing in the definition of semidefinite programs in equation 7 and where we have added to the former vector $\vec{x}$ a real variable $x_0 = \lambda$ which accounts for the fine-tunning of the parameter λ.

To create a single matrix inequality we can direct sum all terms of the inequalities so as to have a large block matrix with all terms of all inequalities.

$$F_i = \bigoplus_{\Omega \subset \mathcal{N}} F_i(\Omega),$$

and in this way we can write the equivalent inequality

$$F_0 + \sum_{i=0}^{d(2^n-n-1)} x_i F_i \geq 0$$

The problem of finding a joint measurability graph, so that the effects will be joint measurable is equivalent to minimizing λ subjected to the constraint above, or subjected to the set of all individual constraints mentioned before In this semidefinite program therefore c will be equal to $c = (1, 0, \dots, 0) \in \mathbb{R}^{d(2^n-n-1)+1}$

$$\min_{\vec{x}} \left\{ c^T \vec{x} = \lambda \,\middle|\, F_0 + \sum_{i=0}^{d(2^n-n-1)} x_i F_i \geq 0 \right\}.$$

- If $\lambda \leq 1$ is the minimum then the effects are joint measurable since the graph can be extended to the unity as shown in figure 4 with a dashed arrow.

- If however $\lambda > 1$ then the effects $\{P_i\}$ are not joint measurable, there exists no joint measure since if it existed we would be able to create a graph such as in figure 4 with $\lambda = 1$, so at least λ should have a value less or equal than 1. *It is worth noticing that the set of operators even if $\lambda > 1$ still build a joint measurability graph, with the only change from theorem 4.1 consisting in considering $S(\emptyset) = \lambda \mathbb{1}$ instead of $\mathbb{1}$.*

The last point will be important from now on, so we find convincing building a proposition out of it:

Proposition 4.1. *To every set of two outcome observables $\{P_1, \dots, P_n\}$ we can assign a joint measurability graph built up out of the solutions for the SDP composed by them.*

Up to now we had seen in theorem 4.1 that we could assign JM-graphs to jointly measurable observables, now we can assign them to any set of effects. In principle we could make them joint measurable by adding noise to the system, like for example considering the convex transformation

$$P_i \to (1 - \eta) P_i + \eta E$$

where $0 \leq E \leq \mathbb{1}$ is some other effect and $\eta \in [0,1]$ (see section 3.2). The problem is of course to know how much noise must be added to the system in order to make the effects joint measurable. In other words, we would like to find constraints for η such that $\alpha \leq \eta \leq \beta$ for certain $\alpha, \beta \in [0,1]$. These constraints should tell us at least how much noise we have to add and if this amount of noise is enough to ensure the joint measurablility of the observables.

In this very sense it has been claimed in [8] that it is possible to define for the case $n = 2$ a distance between the effects accounting for their joint measurablility and amount of noise altogether. Let us therefore consider $\{P_1, P_2\}$ a set of two effects. If $\lambda(P_1, P_2) \leq 1$ these effects would represent the same point, so the distance will have to be 0. Consequently, if we want a distance based on the idea of joint measurability, we represent it by

$$\mu(\{P_i \mid i \in \{1, 2\}\}) = 0, \qquad \text{whenever } \lambda \leq 1$$

If however $\lambda > 1$ then is this extra amount that could be taken in consideration to find a necessary (and also sufficient) noise constraint. This amount of noise is $0 < \mu = 1 - \lambda^{-1} < 1$. Therefore the distance the distance in this case is defined to be:

$$\mu(\{P_i \mid i \in \{1, 2\}\}) = \max\{0, 1 - \lambda^{-1}\}.$$

This was claimed however only for the case of two effects. Regrettably, as it turns out, that the above consideration is inconsistent, because it presupposes that λ is the same performing the SDP with $\{P_1, P_2\}$, $\{P_1, \mathbb{1} - P_2\}$ etc. . . For otherwise there can not be a consistent definition of distance. We we will see that this is the case, λ depends on the combination of POVM elements. We provide later on a correction and a generalization of this fact. For some coming results the idea of joint measurability *subgraphs* will prove to be useful (e.g. Corollary 4.1). We devote next section to their definition and to the prove that they exist.

4.5. **Existence of subgraphs.** In order to use the graph structure to study some aspects about the joint measurability of observables and present them in a graphical and clear way, it has been thought of use to look for substructures in the graph, for example in order to express the joint measurability of a subset of the effects being studied. For that we see that the joint measurability graphs have associated in a straightforward way smaller structures having the same properties and therefore representing subgraphs.

Proposition 4.2. *Let $G = (V, A, S)$ be a n-dimensional joint measurability graph and $\mathcal{N} = \{1, \ldots, n\}$. Then for every $\Omega \subseteq \mathcal{N}$ we can define a* joint measurability subgraph $\tilde{G} = (\tilde{V}, \tilde{A}, \tilde{S})$ *of G in the following way:*

(1) $(\tilde{S} : \Omega \to \tilde{V}) = (S \mid_{\mathcal{N} \cap \Omega} : \mathcal{N} \cap \Omega \to V)$

(2) $\tilde{A} \subseteq A$

Proof. Let us first check the basic properties. From point one we get that $\tilde{S} = S \mid_{\mathcal{N} \cap \Omega}$ and therefore

$$\tilde{V} = \{S(A) \mid A \subseteq \Omega\} \subseteq V.$$

which ensures point 1 in definition 4.1. The second point implies,

$$\tilde{A} = \{(S_1, S_2) \mid ((S_1, S_2) \in A) \wedge (S_1, S_2 \in \tilde{V})\}.$$

which implies points 2 and 4 in the definition of JM-graph. If $\tilde{G}$ is to be a joint measurability subgraph then the condition

$$\tilde{R}(\tilde{\Omega}) \overset{\text{def}}{=} \sum_{A \subseteq \Omega \setminus \tilde{\Omega}} (-1)^{|A|} \tilde{S}(\tilde{\Omega} \cup A) \geq 0$$

must be satisfied for every $\tilde{\Omega} \subset \Omega$. This is indeed the case, to prove it let us make use of the fact that $\tilde{V}$ is a subset of V, then by lemma 4.1 we know that

$$\tilde{S}(\tilde{\Omega} \cup A) = S(\tilde{\Omega} \cup A) = \sum_{A' \subseteq N \setminus (\tilde{\Omega} \cup A)} R(\tilde{\Omega} \cup A \cup A').$$

Let us then rewrite the expression for $\tilde{R}(\tilde{\Omega})$ as

$$\tilde{R}(\tilde{\Omega}) = \sum_{A \subseteq \Omega \setminus \tilde{\Omega}} \sum_{A' \subseteq \mathcal{N} \setminus (\tilde{\Omega} \cup A)} (-1)^{|A|} R(\tilde{\Omega} \cup A \cup A').$$

We need some framework to prove the result. Firstly let us decompose $\mathcal{N} \setminus (\tilde{\Omega} \cup A)$ into two disjoint sets,

$$\mathcal{N} \setminus (\tilde{\Omega} \cup A) = (\mathcal{N} \setminus \Omega) \cup (\Omega \setminus (\tilde{\Omega} \cup A))$$

so since A' is a subset of the first, we can also decompose A' into two subsets, $A' = I \cup O$, which we think as *inside* Ω and *outside* Ω (compare with figure 5 for clarification). We will study the terms of the sum $\tilde{R}(\tilde{\Omega})$ in function of the cardinality of A, where of course $|A| \in \{0, \ldots, \omega - \tilde{\omega}\}$ where $\tilde{\omega} = |\tilde{\Omega}|$ and $\omega = |\Omega|$.
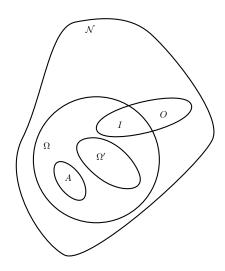

FIGURE 5. Representation of the sets appearing in the proof of proposition 4.2.

Let us take a set $A_k \subset \Omega \backslash \tilde{\Omega}$ with $|A_k| = k$. Then for this set we have in the sum elements of the form

$$R(\tilde{\Omega} \cup A_k \cup I_k \cup O)$$

where $I_k \subseteq \Omega \backslash (\tilde{\Omega} \cup A_k)$. Consider terms using $I_k = \emptyset$ and let us find $A \subset \Omega \backslash \tilde{\Omega}$ and $I \subset \Omega \backslash (\tilde{\Omega} \cup A)$ such that

$$(8) \qquad\qquad R(\tilde{\Omega} \cup A_k \cup \emptyset \cup O) = R(\tilde{\Omega} \cup A \cup I \cup O').$$

Note that if this is the case then $\tilde{\Omega} \cup A_k \cup \emptyset \cup O = \tilde{\Omega} \cup A \cup I \cup O'$ and since $O \cap \Omega = \emptyset$ and $O' \cap \Omega = \emptyset$ it must follow $A_k \cup \emptyset = A \cup I$ and $O' = O$, where we write everywhere $\emptyset$ to express the lack of the I-term in the case of A_k. If we look for candidates on the level of A_k or higher, i.e., we consider A to have k elements or more, the search will fail. Let us suppose that this is indeed the case and there are (A, I) with $|A| \geq k$ such that $A \cup I = A_k$. This can not be the case since $|A \cup I| = |A| + |I| \geq |A_k| = k$, we can only have equality if $|I| = 0$ and $|A| = k$, but then $A = A_k$.

Let us look now for candidates (A, I) with $|A| = j < k$, in this case the equality 8 leads to

$$|A| + |I| = |A_k| \Rightarrow |I| = k - j.$$

Since $A \cup I = A_k$ then A is a subset of A_k of j elements, once the subset A is fixed then I is automatically determined via $A_k \backslash A$. Therefore there are inside the set of A with j elements exactly $\binom{k}{j}$ possibilities of finding a decomposition like $A \cup I = A_k$, moreover the relation

$$A_k \longleftrightarrow \mathcal{C}_j(A_k) \stackrel{\text{def}}{=} \{(A, I) \mid |A| = j, A \cup I = A_k\}$$

is a bijection. Let us sum now all

$$(-1)^{|A|} R(\tilde{\Omega} \cup A \cup I \cup O), \qquad (A, I) \in C_j(A_k)$$

where $j \in \{0, \ldots, k\}$ and $O \subset \mathcal{N} \backslash \Omega$. Since the R-term is the same for all, we will have a coefficient multiplying $R(\tilde{\Omega} \cup A \cup I \cup O)$, let us call this coefficient $c(A_k, O)$, this coefficient will be

$$c(A_k, O) = (-1)^k |C_k(A_k)| + (-1)^{k-1} |C_{k-1}(A_k)| + \ldots + (-1)^0 |C_0(A_k)|$$

and since we said that $|C_j(A_k)| = \binom{k}{j}$ then $c(A_k, O) = 0$ for $k \geq 1$.

The only case where we have $c(A_k, O) \neq 0$ is when $k = 0$, i.e., $A = \emptyset$, in this case we have

$$\mathcal{C}_0(\emptyset) = \{(\emptyset, \emptyset)\}, \quad c(\emptyset, O) = 1$$

and therefore we have

$$\tilde{R}(\tilde{\Omega}) = \sum_{k=0}^{\omega - \tilde{\omega}} \sum_{\substack{A \subset \Omega \backslash \tilde{\Omega} \\ |A| = k}} \sum_{O \subset \mathcal{N} \backslash \Omega} c(A, O) R(\tilde{\Omega} \cup A \cup O) = \sum_{O \subset \mathcal{N} \backslash \Omega} R(\tilde{\Omega} \cup O) \geq 0$$

$\square$

Let us illustrate the result with an example. Suppose $\{P_1, P_2, P_3\}$ are joint measurable and let $\mathcal{N} = \{1, 2, 3\}$. Then by theorem 4.1 we can find a joint measurability graph $G = (V, A, S)$ such that figure 6 holds.

Suppose now $\Omega = \{1, 3\}$ and we want to follow the definition of subgraph to construct it. Following the definition we only need to take the effects P_1 and P_3 and all the elements they have in common up to $S(A)$ where $|A| = 2$. In figure 6 we
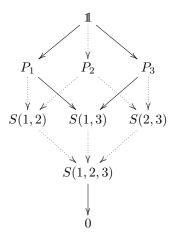
FIGURE 6. Joint measurability subgraph with $n = 3$ and $\Omega = \{1, 3\}$.

have drawn too in solid lines the subgraph. By proposition 4.2 we know that this is a joint measurability subgraph, i.e., we know from the properties of the graph G that $\mathbb{1} - P_1 - P_3 + S(1, 3), P_1 - S(1, 3), P_3 - S(1, 3), S(1, 3) \geq 0$. Physically this means two things:

(1) As it is well known any subset of a set of jointly measurable effects is jointly measurable.

(2) In the graph G for the joint measurability of the effects we have the joint measurability graph for all combinations of the effects.

4.6. **Joint measurability for** $n = 2$**.** Consider the two-outcome observables $\{P, \mathbb{1} - P\}$ and $\{Q, \mathbb{1} - Q\}$, we could build a graph using any combination of elements, $\{P, Q\}$, $\{\mathbb{1} - P, Q\}$ etc. . . Are the parameters coming from the SDP equivalent for every choice of operators? As it may be seen, they are not in general. Take the trivial case where $P = Q = \mathbb{1}$ and perform the SDP using $\{\mathbb{1}, \mathbb{1}\}$ and $\{0, 0\}$. Since they are trivially sharp, the joint measure is unique so that we know exactly the S elements we must write on the graphs. You would have the following representations of the graph:

Of course the infimum value of α so that the graph holds is one and the infimum value of β is 0. Despite this fact, notice that both values are less or equal than 1, which in terms of the information about the joint measurability gives no additional information. However, as a mathematical problem, this last example hints to the fact that in general an invariance of the SDP parameter depending on the choice of POVM elements is not to be expected. We will see through an example that this is indeed too the case for non jointly measurable observables, and we use this fact to discuss a claim provided in [8], which unfortunately we have found to be false.

Claim: [8] Let $\{P, \mathbb{1} - P\}$ and $\{Q, \mathbb{1} - Q\}$ be two observables and let λ_0 be the solution to the SDP constructed by P and Q. Then $\eta = \max\{0, 1 - \lambda_0^{-1}\}$ is the least number such that the two 2-outcome observables induced by the POVM elements $(1-\eta)P + \eta E$ and $(1-\eta)Q + \eta E$ are jointly measurable *for all E* such that $0 \leq E \leq \mathbb{1}$.

This claim is inconsistent with the fact that there might a dependence of the λ SDP parameter on the POVM element choice. Indeed, let us suppose that the parameter λ for (P, Q) is different from the parameter λ' for $(\mathbb{1} - P, \mathbb{1} - Q)$. This

means therefore according to the claim that

$$(1 - \eta)P + \eta E, \qquad (1 - \eta)Q + \eta E, \qquad \forall E(0 \leq E \leq \mathbb{1})$$

are jointly measurable, exactly like

$$(1 - \eta')(\mathbb{1} - P) + \eta'E', \qquad (1 - \eta')(\mathbb{1} - Q) + \eta'E', \qquad \forall E'(0 \leq E' \leq \mathbb{1})$$

are. Here $\eta = \max\{0, 1 - \lambda^{-1}\}$ and accordingly for η'. However that the observables directly above are jointly measurable for every choice of E' means that

$$(1 - \eta')P + \eta'(\mathbb{1} - E'), \qquad (1 - \eta')Q + \eta'(\mathbb{1} - E'), \qquad \forall E'(0 \leq E' \leq \mathbb{1})$$

are jointly measurable by definition. Taking $\mathbb{1} - E' = E$ we have the same condition as before leading to a contradiction, since by hypothesis either $\eta < \eta'$ or $\eta' < \eta$. An example of this behavior, and therefore the proof for this fact, is given in the following example:

Example. Consider $P = \alpha \left|0\right>\!\left<0\right|$ and $Q = \mathbb{1} - \beta \left|\phi\right>\!\left<\phi\right|$ where $\left|\phi\right>$ is such that $\left<0|\phi\right> \neq 0$, $\alpha, \beta \in (0, 1)$ and $\left|\psi\right> = u \left|0\right> + v \left|1\right>$. Of course we are considering qubits with basis $\left|0\right>$ and $\left|1\right>$. To consider the reciprocal SDP, i.e., $\tilde{P} = \mathbb{1} - \alpha \left|0\right>\!\left<0\right|$ and $\tilde{Q} = \beta \left|\phi\right>\!\left<\phi\right|$ notice that the description is exactly the same changing $\alpha \to \beta$ and $\beta \to \alpha$ since $\left|0\right> = u \left|\phi\right> + \tilde{v} \left|\phi^{\perp}\right>$.

Let us therefore consider the SDP with P and Q. From here it is very easy to see which will be the general form of S in the SDP. The condition for P and S is

$$P - S \geq 0 \Rightarrow \alpha \left|0\right>\!\left<0\right| \geq S$$

S can not have any projection on $\left|1\right>$ since $0 \leq \left<1\right| S \left|1\right> \leq \alpha \left<1\right| \left|0\right>\!\left<0\right| \left|1\right> = 0$. Since S is a complex positive semidefinite operator, it must be also hermitian. From these two points we deduce that $S = t \left|0\right>\!\left<0\right|$ where $\alpha - t \geq 0$.

Since $\left|\phi\right> = u \left|0\right> + v \left|1\right>$ and $Q = \mathbb{1} - \beta \left|\phi\right>\!\left<\phi\right|$, we can write Q as

$$Q = \begin{pmatrix} 1 - \beta|u|^2 & -uv^* \\ -u^*v & 1 - \beta|v|^2 \end{pmatrix}$$

For the condition of Q we want

$$S \leq Q = \mathbb{1} - \beta \left|\phi\right>\!\left<\phi\right| \Rightarrow t \left|0\right>\!\left<0\right| + \beta \left|\phi\right>\!\left<\phi\right| \leq \mathbb{1}$$

so the maximum eigenvalue $\lambda_q(t)$ of $t\,|0\rangle\langle 0| + \beta\,|\phi\rangle\langle\phi|$ can be at most 1,

$$\lambda_q(t) = \frac{\sqrt{(t-\beta)^2 + 4\beta t |u|^2} + t + \beta}{2} \leq 1.$$

Since $\lambda_q(t)$ is monotonous increasing, we can find t_1 such that $\lambda_q(t_1) = 1$ and then consider the inequality $t \leq t_1$. It is a matter of checking that

$$t_1(\beta) = \frac{1 - \beta}{1 - \beta + \beta |u|^2}.$$

Therefore we have on the one hand $0 \leq t \leq \alpha$ and on the other $0 \leq t \leq t_1(\beta)$.

The SDP parameter inequality, $\lambda\mathbb{1} - P - Q + S \geq 0$ says that

$$\lambda\mathbb{1} \geq P + Q - S$$

where λ is the infimum of the parameters that fulfill this inequality. Therefore for every t, λ must be the norm of $P + Q - S$. We can write $P + Q - S$ as a matrix,

$$P + Q - S = \begin{pmatrix} 1 - \beta |u|^2 + \alpha - t & -\beta u v^* \\ -\beta u^* v & 1 - \beta |v|^2 \end{pmatrix}$$

The maximum eigenvalue of this matrix (the norm) is

$$\lambda(t; \alpha, \beta, u) = \frac{\alpha - t + 2 - \beta + \sqrt{(\beta + \alpha - t)^2 - 4\beta(\alpha - t)|u|^2}}{2}$$

Therefore this example works by finding an optimal t, which must fulfill $t \leq \alpha$ and $t \leq t_1(\beta)$. Of course if $\alpha \leq t_1(\beta)$ then the optimal choice for t would be α, and otherwise $t_1(\beta)$. On $[0,1]$ $t_1(x)$ is a monotonous decreasing function, and there is no point $x_0 \in (0,1)$ for which $t_1(x_0) = 0$ since $t_1(0) = 1$ and $t_1(1) = 0$.

There exists therefore a point $c \in [0,1]$ for which $c = t_1(c)$, this point is given as a function of u by

$$c(u) = \frac{1}{1 + u}.$$

Therefore, if $\alpha < c(u)$ we have that $\alpha < t_1(x)$ for all $x < c(u)$ (compare with figure 7). This is important since we can choose α and β such that $\alpha \leq t_1(\beta)$ and $\beta \leq t_1(\alpha)$, therefore we would know which is the optimal value of t. This is, in the SDP formed by P and Q, t would be α since $\alpha \leq t_1(\beta)$, we write that as $t_{\alpha\beta} = \alpha$. However, in the SDP formed by $\tilde{P}$ and $\tilde{Q}$, t would be β since $\beta \leq t_1(\alpha)$, we note it as $t_{\beta\alpha} = \beta$. However, in this cases we have

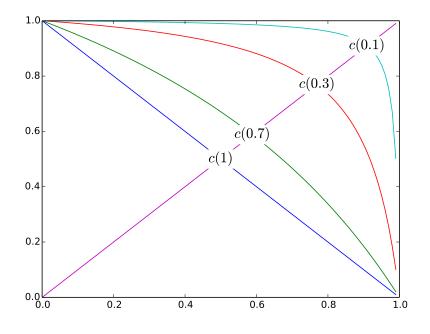$$\lambda_{\alpha\beta} = \lambda(t_{\alpha\beta}; \alpha, \beta, u) = \frac{2 - \beta + \sqrt{\beta^2}}{2} = 1,$$

FIGURE 7. Representation of the functions $x \mapsto x$ and $t_1(x; u)$ for $u \in \{1, 0.7, 0.3, 0.1\}$. The intersection point $c(u)$ has been drawn too for each u.

and the same for $\lambda_{\beta\alpha}$. Therefore, the set of points

$$M_{<<} = \{(\alpha, \beta) \mid \alpha < t_1(\beta), \beta < t_1(\beta)\}$$

lead to a SDP parameter invariant (from P, Q to $\tilde{P}, \tilde{Q}$) and equal to 1. It is clear that whenever $\alpha = \beta$, the equation for $\lambda(t, \alpha, \beta, u)$ is symmetrical and we have invariance too. There are two sets left to consider, i.e.

$$M_{<>} = \{(\alpha, \beta) \mid \alpha < t_1(\beta), \beta > t_1(\alpha), \alpha \neq \beta\}$$

and

$$M_{>>} = \{(\alpha, \beta) \mid \alpha > t_1(\beta), \beta > t_1(\alpha), \alpha \neq \beta\}.$$

For these two sets we make two conjectures. The first is that $M_{<>} = \emptyset$. The second is that the set of points for which $\lambda(t, \alpha, \beta, u)$ is not invariant under the change $\alpha \leftrightarrow \beta$ is $M_{>>}$. These conjectures are backed by numerical calculations performed using the algorithm found in the appendix A.

Some points where the SDP parameters are not invariant are presented in table 1.

TABLE 1. Some suitable points where the SDP parameters $\lambda_{\alpha\beta}$ and $\lambda_{\beta\alpha}$ are not equal.

α	β	$t_1(\alpha)$	$t_2(\beta)$	u	$\lambda_{\alpha\beta}$	$\lambda_{\beta\alpha}$	$\in M_{>>}$?
1.0	0.2	0.0	0.83	0.9	1.06574	1.04497	yes
0.9	0.7	0.18	0.47	0.7	1.28205	1.33054	yes
0.8	0.9	0.28	0.15	0.8	1.3521	1.34139	yes
0.7	0.6	0.54	0.65	0.6	1.03338	1.0372	yes
0.5	0.9	0.74	0.24	0.6	1.18529	1.11663	yes
0.4	0.9	0.86	0.31	0.5	1.07092	1.03296	yes
0.3	0.9	0.87	0.24	0.6	1.04209	1.02238	yes
0.2	0.9	0.86	0.15	0.8	1.01945	1.01539	yes

Refinement. A suitable refinement of the *claim* we have found to be the following proposition:

Proposition 4.3. *Let $\{P_1, \mathbb{1} - P_1\}$ and $\{P_2, \mathbb{1} - P_2\}$ be two observables. Consider the SDP parameter λ_0 coming from considering the SDP for $\{P_1, P_2\}$. Then the parameter η given by*
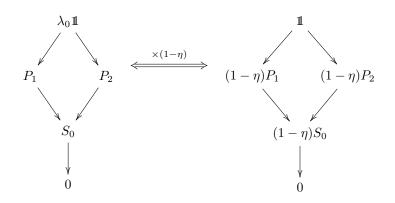
$$\eta = \max\{0, 1 - \lambda_0^{-1}\}$$

is the least number such that $\{(1-\eta)P_1, (1-\eta)(\mathbb{1} - P_1) + \eta\mathbb{1}\}$ and $\{(1-\eta)P_2, (1-\eta)(\mathbb{1} - P_2) + \eta\mathbb{1}\}$ are jointly measurable. Furthermore, if

$$\lambda_0\mathbb{1} - P_1 - P_2 + S_0 \neq 0$$

where S_0 is a solution to the SDP, there exist some operators E with $0 < E < \mathbb{1}$ such that the observables $\{(1-\eta)P_i+\eta E, (1-\eta)P_i+\eta(\mathbb{1}-E)\}$ are jointly measurables for $i \in \{1,2\}$.

Proof. Let us suppose that $\lambda_0 > 1$, for otherwise the claim is automatic since the observables are jointly measurable and $\eta = 0$. The structure of JM-graphs is invariant under multiplication by a positive scalar. Indeed, for every element we have an inequality the sign of which is not changed by the product. Therefore

taking $1 - \eta = \lambda_0^{-1} \in (0, 1)$ we have the following representation:

For the SDP composed by $(1-\eta)P_1$ and $(1-\eta)P_2$, 1 is the minimum SDP parameter, since λ_0 was the minimum for P_1 and P_2. Now consider the SDP for $(1-\eta)P_i + \eta E$,

$$\mathbb{1} - (1 - \eta)P_1 - (1 - \eta)P_2 - 2\eta E + S \geq 0$$
$$(1 - \eta)P_i + \eta E - S \geq 0$$
$$S \geq 0$$

If we prove the feasibility of this system for some S and E then they will be jointly measurable. Indeed take $S = (1 - \eta)S_0$, and E such that

(9) $$\mathbb{1} \geq^3 \mathbb{1} - (1 - \eta)P_1 - (1 - \eta)P_2 + (1 - \eta)S_0 \geq 2\eta E \geq 0$$

If $\lambda_0\mathbb{1} - P_1 - P_2 + S_0 \neq 0$ equation 9 allow us to find some $0 < E < \mathbb{1}$ (*near* 0) such that with this choice of E and S the system's feasibility is fulfilled, i.e.:

$$\mathbb{1} - (1 - \eta)P_1 - (1 - \eta)P_2 - 2\eta E + (1 - \eta)S_0 \geq 0$$
$$(1 - \eta)P_i - (1 - \eta)S_0 + \eta E \geq (1 - \eta)P_i - (1 - \eta)S_0 \geq 0$$
$$S_0 \geq 0$$

$\square$

This proposition is a first step towards a full characterization of the problem. Firstly, the parameter η is still dependent on the choice of the POVM elements to build the SDP or the JM-graph with. Therefore, it does not seem to be possible to consider η as a *full* descriptor of the noise of the joint measurement of $\{P_1, \mathbb{1} - P_1\}$ and $\{P_2, \mathbb{1} - P_2\}$ as claimed in [8]. We have sought to resolve this problem in proposition 4.4, where we have found that such a parameter with the desired interpretation is obtainable.

---

[3]By the properties of S_0 we know that $-(1-\eta)P_1 + (1-\eta)S_0 \leq 0$, so $\mathbb{1} - (1-\eta)P_2 - (1-\eta)P_1 + S \leq \mathbb{1} - (1 - \eta)P_2 \leq \mathbb{1}$ since $P_2 \geq 0$.

However, before setting out to prove the proposition we may ask about the possibility that depending on the choice of the POVM elements some SDP parameters be less or equal than one and some greater. More explicitly:

Lemma 4.2. *Let $\{P_1, P_2\}$ and $\{Q_1, Q_2\}$ bet two. Let $\lambda(i,j)$ be the solution to the SDP problem taking P_i and Q_j. If there exist $i_0, j_0 \in \{1, 2\}$ such that $\lambda(i_0, j_0) \leq 1$, then $\lambda(i,j) \leq 1$ for all $i, j \in \{1, 2\}$.*

Proof. Suppose there exist such (i_0, j_0) as in the hypothesis. Then $\{P_1, P_2\}$ and $\{Q_1, Q_2\}$ are jointly measurable, which means that we can find a joint measure R and build graphs as in theorem 4.1, which for any choice of P_i and Q_j would give us a feasibility condition for the SDP parameter $\lambda = 1$, therefore $\lambda(i,j) \leq 1$. $\square$

Proposition 4.4. *Let $\{P_1, P_2\}$, $\{Q_1, Q_2\}$ and $\lambda(i,j)$ be as in lemma 4.2.*

(1) *There exists a minimum parameter $\eta \in [0, 1]$ such that $\{(1-\eta)P_i, \mathbb{1} - (1-\eta)P_i\}$ and $\{(1-\eta)Q_j, \mathbb{1} - (1-\eta)Q_j\}$ are jointly measurable for any choice of $i, j \in \{1, 2\}$. This parameter is given by $\max\{0, 1 - \lambda_*^{-1}\}$ where*

$$\lambda_* = \max\{\lambda(i,j) \mid i, j \in \{1, 2\}\}$$

Proof.   (1) If $\lambda_* \leq 1$ then the proposition holds since they are jointly measurable and therefore $\eta = 0$. Suppose then $\lambda_* > 1$, then by lemma 4.2 we have $\lambda(i,j) > 1$ for all $i, j \in \{1, 2\}$ and conditions for proposition 4.3 are met. Therefore considering η as defined in the hypothesis would grant us the joint measurability of $(1-\eta)P_i$ and $(1-\eta)Q_j$ for all, $i, j \in \{1, 2\}$ since η is by definition greater or equal than the minimum parameter η' in each case.
$\square$

We leave as an open problem to show under exactly which conditions the SDP parameter is invariant depending on the choice of the POVM elements. Also we find interesting to characterize the set of feasible E's, i.e., for which $0 \leq E \leq \mathbb{1}$ exactly the convex combinations are jointly measurable.

4.7. **Joint measurablility via subgraphs.** Using the idea of joint measurability graphs we have found a straightforward way of defining information parameters that account for the joint measurablility of subsets of two outcome observables and the j-wise joint measurability of n two outcome observables. These parameters have to do with the minimum amount of noise required to make the subset of effects joint measurable in the same way as in the last subsection.

Let $\{P_1, \ldots, P_n\}$ be n two outcome observables and $G = (V, A, S)$ the joint measurability graph associated to the set of effects *through the semidefinite program*, i.e., assuming $S(\emptyset) = \lambda_{\min}\mathbb{1}$ (see subsection 4.4 and proposition 4.1). Let $\Omega \subseteq \mathcal{N}$ with $|\Omega| \geq 2$ and $G(\Omega)$ be the subgraph induced by Ω as in proposition 4.2.

For this subset Ω we can write the semidefinite program as in section 4.4 to obtain a parameter value λ which we will write in dependence of Ω as $\lambda(\Omega)$. Let us state a property about the semidefinite program for the subgraph $G(\Omega)$. As we know it is about looking for the $\inf\limits_{\{S(A), A \subset \Omega, |A| \geq 2\}} \{\alpha\}$ subjected to

$$(10) \quad \begin{cases} \alpha\mathbb{1} + \sum\limits_{\substack{A \subset \Omega \\ A \neq \emptyset}} (-1)^{|A|} S(A) \geq 0 \\ \\ \sum\limits_{A \subseteq \Omega \setminus \tilde{\Omega}} (-1)^{|A|} S(\tilde{\Omega} \cup A) \geq 0 \quad \text{for all } \tilde{\Omega} \subset \Omega \text{ where } \tilde{\Omega} \neq \emptyset \end{cases}$$

The fact that $\lambda(\Omega)$ is the infimum means that the constraints are being fulfilled for $\alpha \geq \lambda(\Omega)$. If however we consider an $\alpha_0 < \lambda(\Omega)$, then necessary the first equation in 10 must cease to be valid since the other ones do not depend on the parameter α and will continue to hold regardless of it. In other words, the semidefinite positivity of the first equation would contradict the nature of $\lambda(\Omega)$. Note however that this does not mean that the first inequality must be negative, since the order relation $\geq$ is not a full order for operators acting on a Hilbert space. As stressed for the case of $n = 2$, the parameter $\lambda(\Omega)$ will in general depend on the choice of POVM element for every observable, i.e., depending on whether or not we choose P_i or $\mathbb{1} - P_i$, the name of which is arbitrary. For the last section, the conclusion was to consider the maximum of these parameters in order not to have this problem, i.e., $\lambda^*(\Omega) = \max \lambda(\Omega)$. We will do here the same thing, however, since we will use always λ^* for its interpretation, we can rename it as λ. We think it is important to formalize this concept.

Definition 4.2. *Let us consider n observables $\mathsf{P}_i = \{P_i^1, P_i^2\}$ where $i \in \mathcal{N} = \{1, \ldots, n\}$.[4] For every choice of POVM element from the observables, i.e, for every set $\{P_1^{i_1}, \ldots, P_n^{i_n}\}$ where all $i_k \in \{1, 2\}$ we can find the SDP parameter $\lambda(i_1, \ldots, i_n)$ formed by the POVM elements. Let us define by $\lambda(\mathcal{N})$ the maximum of these SDP parameters considering all choices, i.e.*

$$\lambda(\mathcal{N}) = \max\{\lambda(i_1, \ldots, i_n) \mid i_k \in \{1, 2\}, k \in \mathcal{N}\}$$

With this definition we can therefore define a parameter $\lambda_i(\mathcal{N})$ for every storey, i.e., for every subset of $|\Omega|$ elements. This parameter will be analogous to the parameters for the case $n = 2$, but taking into consideration the j-wise measurability.

Definition 4.3. *Let $i \in \{2, \ldots, n\}$ and P_i be as in definition 4.2. For every $\Omega \subseteq \mathcal{N}$, the parameter $\lambda(\Omega)$ at the beginning of this section, where we consider the set of observables $\{\mathsf{P}_i \mid i \in \Omega\}$ for the SDP. Therefore we can define $\lambda_i(\mathcal{N})$ as*

$$\lambda_i(\mathcal{N}) = \max\{\lambda(\Omega) \mid \Omega \subseteq \mathcal{N}, |\Omega| = i\}.$$

As it can be seen from the definition, each parameter (which remember are effectively obtainable) $\lambda^i(\mathcal{N})$ accounts for the joint measurability of the effects i-wisely, in the following sense:

- If $\lambda^i(\mathcal{N}) \leq 1$ then the observables $\{\mathsf{P}_i \mid i \in \mathcal{N}\}$ are joint measurable i-wisely, since $\lambda(\Omega) \leq \lambda^i(\mathcal{N})$ for every Ω with $|\Omega| = i$.
- If however $\lambda^i(\mathcal{N}) > 1$ then it is the negation of the sentence above since at least there exists one $\Omega \subset \mathcal{N}$ with $|\Omega| = i$ so that $\lambda(\Omega) > 1$. Therefore they would not be i-wisely measurable.

We have found that these $\lambda^i(\mathcal{N})$ are ordered in a regular way inversely following the ordered structure of the storeys in the graph, therefore it is helpful to state it down precisely in the following proposition:

Proposition 4.5. *Let G and $\mathcal{N}$ be as at the beginning of the section. The following inequality is satisfied*

$$\lambda^i(\mathcal{N}) \leq \lambda^{i+1}(\mathcal{N})$$

for $i \in \{2, \ldots, n-1\}$.

Proof. Let $\Omega_0 \subset \mathcal{N}$ be such that $|\Omega_0| = i$ and $\lambda^i(\mathcal{N}) = \lambda(\Omega_0)$. Take any set $\Omega \subseteq \mathcal{N}$ such that $\Omega_0 \subset \Omega$ and $|\Omega| = i + 1$ which exists since $i < n$. We will prove the result

---

[4] To avoid confusions let us state that $P_i^1 + P_i^2 = \mathbb{1}$ for all $i \in \mathcal{N}$.

supposing that $\lambda(\Omega) < \lambda(\Omega_0)$ and getting to a contradiction, therefore $\lambda^{i+1}(\mathcal{N})$ will be at least as big as $\lambda(\Omega_0)$. The infimum condition for $\alpha = \lambda(\Omega)$ in equation 10 means that one can build a joint measurability graph with $S(\emptyset) = \lambda(\Omega)\mathbb{1}$ (of course we should select a suitable choice of the POVM elements). As a joint measurability graph, $G(\Omega)$ contains several subgraphs, in particular the subgraph generated by Ω_0. This subgraph has $S(\emptyset) = \lambda(\Omega)$ and it follows therefore that

$$
\begin{cases}
\lambda(\Omega)\mathbb{1} + \displaystyle\sum_{\substack{A \subset \Omega_0 \\ A \neq \emptyset}} (-1)^{|A|} S(A) \geq 0 \\[2em]
\displaystyle\sum_{A \subseteq \Omega_0 \setminus \tilde{\Omega}} (-1)^{|A|} S(\tilde{\Omega} \cup A) \geq 0 \quad \text{for all } \tilde{\Omega} \subset \Omega_0 \text{ where } \tilde{\Omega} \neq \emptyset
\end{cases}
$$

which can not be since $\lambda(\Omega) < \lambda(\Omega_0)$ and $\lambda(\Omega_0)$ is the least number for which this system of inequalities would hold. Therefore $\lambda(\Omega) \geq \lambda(\Omega_0) = \lambda^i(\mathcal{N})$. $\qquad\square$

Physically, proposition 4.5 accounts for the well-known fact that if a set of n obersvables is not i-wisely joint measurable, then it can not be either $(i+1)$-wisely joint measurable. It also accounts for the fact that we can not make a set of observables jointly measurable by adding more obsevables.

Proposition 4.5 is also interesting on its own for the following question: is it possible to have a set of observables $\{\mathsf{P}_i \mid i \in \mathcal{N}\}$ for which some parameters $\lambda_j(\mathcal{N}) < \lambda_{j+1}(\mathcal{N})$? This would mean for example, in the case that there existed a $j \in \{2, \ldots, n-1\}$ such that $\lambda_j(\mathcal{N}) \leq 1$ and $\lambda_{j+1}(\mathcal{N}) > 1$, the observables would be j-wise measurable but not $(j+1)$-wise, which is a treat characteristic of POVM's in general, as opposed to sharp POVM's. Actually we can prove that such an ordering is possible. In the recent paper of Kunjwal et al. [3], they have shown that all jointly measurable structures are quantum realizable. This meaning that all possible scenarios regarding measurability are attainable using observables. For example there exists a set of n observables $\{\mathsf{A}_i\}$ such that they are j-wise jointly measurable but not $(j+1)$-wise, for any $j < n$ we may fancy. In fact, they provide the existence using *two-outcome observables*, which ensures that in our structure we can always find some set $\{\mathsf{P}_i\}$ of two-outcome observables such that we have $\lambda_j(\mathcal{N}) < \lambda_{j+1}(\mathcal{N})$ for some desired j. We put this result in form of a proposition to lay it down.

Proposition 4.6. *For every $i > 2$ there exists some set of n two-outcome observables, where $n > i$, for which in the notation of proposition 4.5*

$$\lambda_i(\mathcal{N}) < \lambda_{i+1}(\mathcal{N}).$$

Up to now however $\lambda^i(\mathcal{N})$ have been mere parameters which have something to say about the joint measurablility of subsets. What follows makes the case for their importance in constraining and fixing the amount of noise that must be added into a system to make the effects jointly measurable. The following two propositions must be seen as a generalization of the results on section 4.6, where we found corrections and refinements of the results in [8]. We put together in proposition 4.7 the generalization of propositions 4.3 and 4.4.

Proposition 4.7. *Let $\{\mathsf{P}_i \mid i \in \mathcal{N} = \{1, \ldots, n\}\}$, with $n \geq 2$ be a set of 2-outcome observables, where $\mathsf{P}_i = \{P_i^1, P_i^2\}$.*

(1) *Let $\lambda(i_1, \ldots, i_n)$ and S_0 be the solutions to the SDP taking into consideration the POVM elements $\{P_1^{i_1}, \ldots, P_1^{i_n}\}$ with $i_k \in \{1, 2\}$ for all $k \in \mathcal{N}$. Then the parameter*

$$\eta = \max\{0, 1 - \lambda(i_1, \ldots, i_n)^{-1}\}$$

is the least number such that the effects $\{(1-\eta)P_j^{i_j}, \mathbb{1} - (1-\eta)P_j^{i_j}\}$ are jointly measurable for all $j \in \mathcal{N}$. Furthermore, if

(11)
$$\lambda(i_1, \ldots, i_n)\mathbb{1} - \sum_{j=1}^n P_j^{i_j} + \sum_{\substack{A \subset \mathcal{N} \\ |A| \geq 2}} S_0(A) \neq 0$$

then there exists E, $0 < E < \mathbb{1}$ such that $\{(1-\eta)P_j^{i_j} + \eta E, \mathbb{1} - (1-\eta)P_j^{i_j} - \eta E\}$ are jointly measurable for all $j \in \mathcal{N}$.

(2) *There exists a parameter η such that $\{(1-\eta)P_j^{i_j}, \mathbb{1} - (1-\eta)P_j^{i_j}\}$ are jointly measurable for all $j \in \mathcal{N}$ regardless of the specific choice of i_j. Therefore one can write for this parameter that*

$$\{(1-\eta)\mathsf{P}_j \mid j \in \mathcal{N}\}$$

are jointly measurable. This parameter is given by

$$\eta = \max\{0, 1 - \lambda(\mathcal{N})^{-1}\}$$

where $\lambda(\mathcal{N})$ is defined in definition 4.2.

Proof. (1) The arguments for the first claim are exactly the same than for the case in proposition 4.3. The case for $\lambda(i_1,\ldots,i_n) \leq 1$ is clear, let us suppose otherwise. One is able to write the inequalities using the notation of JM-graphs, i.e.,

$$\lambda(i_1,\ldots,i_n)\mathbb{1} - \sum_{j\in\mathcal{N}} P_j^{i_j} + \sum_{\substack{A\subset\mathcal{N} \\ |A|\geq 2}} (-1)^{|A|}S_0(A) \geq 0$$

$$P_j^{i_j} + \sum_{\substack{A\subset\mathcal{N}\setminus\{j\} \\ |A|\neq\emptyset}} (-1)^{|A|}S_0(\{i\}\cup A) \geq 0, \quad \forall j\in\mathcal{N}$$

$$\sum_{A\subset\mathcal{N}\setminus\Omega} (-1)^{|A|}S_0(\Omega\cup A) \geq 0$$

where $\lambda(i_1,\ldots,i_n)$ is the infimum. Multiplying these inequalities by some positive constant k does not change the infimum nature of the SDP parameter $k\lambda(i_1,\ldots,i_n)$, in particular if $k = \lambda(i_1,\ldots,i_n)^{-1} = 1 - \eta$ then the infimum becomes one, so it is the minimum for which we have joint measurability of $(1-\eta)P_j^{i_j}$. Next, if we consider the SDP for some E and feasibility for $S = (1-\eta)S_0$, we get considering equation 11

$$\overbrace{\mathbb{1} - \sum_{j\in\mathcal{N}}(1-\eta)P_j^{i_j} + \sum_{\substack{A\subset\mathcal{N} \\ |A|\geq 2}} (-1)^{|A|}(1-\eta)S_0(A)}^{>0} - n\eta E \geq 0$$

$$\eta E + \overbrace{(1-\eta)P_j^{i_j} + \sum_{\substack{A\subset\mathcal{N}\setminus\{j\} \\ |A|\neq\emptyset}} (-1)^{|A|}(1-\eta)S_0(\{i\}\cup A)}^{\geq 0} \geq 0, \quad \forall j\in\mathcal{N}$$

$$\sum_{A\subset\mathcal{N}\setminus\Omega} (-1)^{|A|}(1-\eta)S_0(\Omega\cup A) \geq 0$$

some values of E, where $0 < E < \mathbb{1}$ for which the SDP is feasible.

(2) This is also analogous to the proof of proposition 4.4. By definition,

$$\lambda(\mathcal{N}) = \max\{\lambda(i_1,\ldots,i_n) \mid i_k \in \{1,2\}, k\in\mathcal{N}\}.$$

For any choice of i_k, $\{(1-\eta)P_j^{i_j}\}$ will be jointly measurable since the minimum parameter $\eta(i_1,\ldots,i_n)$ for the given choice is less than η. This is due to the fact that $\lambda(i_1,\ldots,i_n) \leq \lambda(\mathcal{N})$ and that the function $x \rightarrow 1-x^{-1}$ is monotonous increasing for $x > 0$.

$\square$

Corollary 4.1 (*j*-wise joint measurablility)**.** *Let $\{\mathsf{P}_i \mid i \in \mathcal{N}\}$ be a set of n 2-outcome observables and $j \in \{2, \ldots, n\}$. The minimum amount of noise η_j needed to make the observables $(1 - \eta_j)\mathsf{P}_i$ j-wisely measurable is given by*

$$\eta_j = \max\{0, 1 - \lambda_j(\mathcal{N})^{-1}\}$$

where $\lambda_j(\mathcal{N})$ was defined in definition 4.3 on page 35.

Proof. For every $\Omega \subset \mathcal{N}$ where $|\Omega| = j$ consider the subgraph $G(\Omega)$. To each subgraph we have the parameter $\lambda(\Omega)$ associated by the SDP. As a graph there is for every $G(\Omega)$ a noise parameter $\eta(\Omega)$ from proposition 4.7 which is the least amount of noise necessary to make $\{(1 - \eta(\Omega)P_i \mid i \in \Omega\}$ joint measurable (notice that we need not consider the choice of the POVM elements P_i due to the definition of $\lambda(\Omega)$). Simply take the maximum of these $\eta(\Omega)$ as η_j. $\qquad\square$

4.8. **Universality of noise.** In this section we will use some results about last section and use them to prove the property that there exist noise parameters η for which all effects are k-wisely measurable, we give constraints and actual values for η.

Lemma 4.3. *Let $\mathcal{N} = \{1, \ldots, n\}$ and $\{P_1, \ldots, P_n\}$ 2-outcome observables. Then*

$$\lambda(\mathcal{N}) \leq n$$

and as a consequence

$$\lambda^i(\mathcal{N}) \leq i.$$

Proof. Let $\lambda(\mathcal{N})$ and S be an optimal solution of the SDP for $\{P_i\}$. The first inequality gives us

$$\lambda(\mathcal{N})\mathbb{1} - \sum_{i \in \mathcal{N}} P_i + \sum_{\substack{A \subset \Omega \\ |A| \geq 2}} (-1)^{|A|} S(A) \geq 0$$

As a mathematical problem however, the feasible convex set of this convex optimization problem includes the point $(n, \{S(\Omega) = 0 \mid \Omega \subset \mathcal{N}, |\Omega| \geq 2\})$ since for all $i \in \mathcal{N}$ we have $0 \leq P_i \leq \mathbb{1}$, therefore

$$n\mathbb{1} - \sum_{i \in \mathcal{N}} P_i \geq 0$$

and the other constraints are met *trivially*. Therefore the optimal solution parameter $\lambda(\mathcal{N})$ must be less or equal than n. The condition for $\lambda^i(\mathcal{N})$ is a consequence of the general one considering JM-subgraphs. $\qquad\square$

This last lemma is useful to prove the following proposition, which get to the idea that noise can be made universal for all 2-outcome observables.

Proposition 4.8 (Universality of noise). *Let $j \geq 2$. There exists a parameter $\eta_j \in [0, 1 - j^{-1}]$ such that for every set 2-outcome observables $\{\mathsf{P}_i \mid i \in \mathcal{N} = \{1. \ldots, n\}\}$ with $n \geq 2$, the convex transformed observables $\{(1 - \eta^j)\mathsf{P}_i \mid i \in \mathcal{N}\}$ are j-wisely jointly measurable.*

Proof. Consider $\tilde{\eta}_j = 1 - \lambda_j(\mathcal{N})^{-1}$. We know from corollary 4.1 that for any $\eta \geq \tilde{\eta}_j$ the set of effects

$$\{(1 - \eta)\mathsf{P}_i \mid i \in \mathcal{N}\}$$

is j-wise jointly measurable. However from lemma 4.3 we know too that independently of $\{\mathsf{P}_i\}$

$$\lambda_j(\mathcal{N}) \leq j$$

therefore we may as well choose a η independent of $\mathcal{N}$ since $1 - \lambda^j(\mathcal{N}) \leq 1 - j^{-1}$. Therefore we are always able to find such an η in $[0, 1 - j^{-1}]$, where in particular $\eta = 1 - j^{-1}$ will work for any set of effects, but it is not ensured to be the minimum.

$\square$

Despite proposition 4.8 we should not be led to think that actually we can find a noise parameter so that all effects are *all-wise* jointly measurable. For this case proposition 4.8 should be addressed using a limiting process. We could perform a limiting process taking a set $\mathcal{N}(n) = \{1, \ldots, n\}$ and n effects $\{\mathsf{P}_i\}_{i \in \mathcal{N}(n)}$. We should increase the size of the set, thus watching what happens with the parameter η_n. Of course as n tends to infinity the closed interval $[0, 1 - n^{-1}]$ tends to $[0, 1]$, which does not ensure that $\eta_\infty < 1$. It is therefore left as an open question whether or not $\eta_\infty < 1$. In the same way it is left to show whether or not $\eta_j = 1 - j^{-1}$ or $\eta_j < 1 - j^{-1}$. The case $\eta_j = 1 - j^{-1}$ seems more appealing as the limiting process would get us $\eta_\infty = 1$, however it is important to note that proposition 4.8 only specifies that $\eta_j \in [0, 1 - j^{-1}]$.

4.9. **Universality of noise for general POVM's.** In this section we will provide an immediate generalization of last results for a particular case of general multiple outcome observables, where at least one of them is assumed to be sharp. The main idea is that from every general observable we can always produce several two outcome observables. The question arises, how can we relate the joint measurability of these "subobservables" or "partitions" to the joint measurability of the *bigger* observables? For that we need a definition and we borrow a result that allows us to generalize proposition 4.8.

Definition 4.4. *Let C be a POVM, Ω_C its outcome space and $\mathcal{F}_C$ its σ-algebra. Take $X \in \mathcal{F}_C$, then the effect $\{C(X), \mathbb{1} - C(X)\}$ is called a partition of C or a X-partition of C. Physically, it is the POVM that is to measure if a given outcome belongs to X.*

Proposition 4.9 (From [2]). *Suppose at least A or B (or both) is sharp. Then they are jointly measurable if and only if every partition of A and B is jointly measurable.*

Last proposition can seem natural at first glance, however as it is noted in [2], the assumption that at least one of them is sharp is *essential*. This means, if neither A nor B are sharp, the result ceases to be valid for some observables. Let us present our use of proposition 4.9.

Proposition 4.10. *Let A and B be POVM where at least one of them is sharp. Then there exists an $\eta \in \left[0, \frac{1}{2}\right]$ such that the convex compositions $(1 - \eta)A$ and $(1 - \eta)B$ are jointly measurable.*

Proof. Let $(\Omega_A, \mathcal{F}_A)$ and $(\Omega_B, \mathcal{F}_B)$ be the outcome spaces of A and B respectively and take $X \in \mathcal{F}_A$, $Y \in \mathcal{F}_B$. Let us consider the partitions $A^X = \{A(X), \mathbb{1} - A(X)\}$ and $B^Y = \{B(Y), \mathbb{1} - B(Y)\}$. We can consider as throughout the section $\mathsf{P}_1 = A^X$ and $\mathsf{P}_2 = B^Y$ two observables, for which a parameter $\lambda(X, Y)$ can be calculated that makes $(1 - \lambda(X, Y)^{-1})\mathsf{P}_i$ joint measurable for $i \in \{1, 2\}$. Let us therefore consider

$$\lambda = \sup_{\substack{X \in \mathcal{F}_A \\ Y \in \mathcal{F}_B}} \{\lambda(X, Y)\},$$

which also fulfills $\lambda \leq 2$ by virtue of lemma 4.3. Taking $\eta = 1 - \lambda^{-1}$, we know that $(1 - \eta)A^X$ and $(1 - \eta)B^Y$ are joint measurable for whatever choice of $X \in \mathcal{F}_A$ and $Y \in \mathcal{F}_B$. Therefore by proposition 4.9 $(1 - \eta)A$ and $(1 - \eta)B$ are joint measurable and $\eta \in \left[0, \frac{1}{2}\right]$. $\square$

The above proposition is completely general inasmuch that it does not consider discrete observables at all, i.e, We may therefore consider continuous observables.

4.10. **Definition of the distance.** We are able at this point to generalize the notion of distance proposed in [8], where a "distance" for the case of two effects was proposed (see at the end of subsection 4.4). We propose two different definitions, which rectify the idea proposed in [8].

Definition 4.5. *Let $\mathcal{N} = \{1, \ldots, n\} \subset \mathbb{N}$ and let $\{P_i \mid i \in \mathcal{N}\}$ be any set of n effects. We can define a "distance" μ as*

$$\mu(P_1, \ldots, P_n) = \eta(\mathcal{N})$$

where $\eta(\mathcal{N}) = \max\left\{0, 1 - \lambda(\mathcal{N})^{-1}\right\}$.

The following definition takes into account too the parameters which are related to the subgraphs or to j-wise measurability, where $j \leq n$.

Definition 4.6 (With subgraphs)**.** *Let $\mathcal{N} = \{1, \ldots, n\} \subset \mathbb{N}$ and let $\{P_i \mid i \in \mathcal{N}\}$ be any set of n effects. We can define a "distance" μ_s[5] as*

$$\mu_s(P_1, \ldots, P_n) = \sum_{j=2}^{n} \eta^j(\mathcal{N})$$

where $\eta^j(\mathcal{N}) = \max\left\{0, 1 - \lambda^j(\mathcal{N})^{-1}\right\}$.

From the definitions it is worth mentioning that whenever $\mu(P_1, \ldots, P_n) = 0$ or $\mu_s(P_1, \ldots, P_n) = 0$, it means that the observables are jointly n-wisely measurable and they would therefore represent the same point. Note that these definitions are consistent since $\lambda^k(\mathcal{N})$ is independent of the choice of POVM elements, unlike the distance proposed in [8].

---

[5]We write μ_s to denote that it is built upon subgraphs.

5. Conclusion

The main finding of this work is the following theorem:

Theorem 5.1. *To every set of two outcome observables $\{\mathsf{P}_i \mid i \in \mathcal{N} = \{1, \ldots, n\}\}$ we can associate a joint measurability graph $G = (V, A, S)$ coming as a solution of a semi definite program, together with efficiently computable parameters $\eta^j(\mathcal{N})$ for every storey, which account for the minimum amount of noise needed to make the observables $\{(1 - \eta^j)\mathsf{P}_i \mid i \in \mathcal{N}\}$ j-wise jointly measurable.*

For sharp observables there is a complete operational characterization of joint measurability, i.e., in that case it is equivalent to commutativity. However, for general observables, this is not the case anymore, as we have already stressed throughout the report. Therefore, this theorem could open the door to general methods using homology theory to describe the joint measurability of observables. The direction to follow could be to define an algebraic structure within the set of joint measurability graphs, so that one may be able to add and substract them in order to obtain particular or general results. This structure in turn could give us some insight about a full characterization of the joint measurability of two outcome observables, and maybe of general ones. For that we think the joint measurability graphs could be useful.

There are several treats of this work from which other people looking at these problems could benefit, namely: the use of the notation of joint measurability graphs and the generalizations that this notation enables for the case of n observables. Also the graph structure and its notation can help to devise general algorithms to perform SDP for a large number of two outcome observables.

However, no report is complete without a critique of the ideas and the results obtained. It is not yet clear whether this graph structure is going to be useful or essential to future research in the field. In our opinion, the main reason to be doubtful is the dependence of the SDP parameter on the choice of POVM elements to perform the SDP with. The SDP parameters do not seem to bear anymore the whole information about the joint measurability due to this fact, that is why the discovery of the mistake in [8] has been crucial. In consequence, the distances proposed should be studied in their own generality to asses their relevance considering the lack of invariance in the SDP parameters. This problem could in turn be quite serious to whether or not considering SDP's to study the quantification of joint measurability for observables. It is an open matter whether we should turn our

attention to noise operators E in convex combinations rather than noise parameters η, or maybe to a combination of both, to have a consistent interpretation of the information obtained through SDP's. Although we have given an approach to mending this hindrance by taking the maximum of these parameters, it does not seem to be an elegant solution to describe the noise of the systems. Also, as the dimension of the problem rises, the number of combinations needed to calculate the SDP parameters (definition 4.2) rises exponentially, conforming thus a threat to the effective computability of the parameters. More work should be done to clarify these terms and to establish a complete characterization of the existence of POVM's E in noisy convex combinations $(1-\eta)\mathsf{P}_i + \eta\mathsf{E}$. This could ultimately lead to a complete description of this kind of noisy models for two outcome observables and maybe even for general observables. In this area our results have been modest. Indeed, we only consider two general observables and we had to assume that one of them is sharp.

We hope to have shed some light on the named aspects about the joint measurability of two outcome observables.

Appendix A. Program to look for counter-examples

The following algorithm has the following structure:

(1) Generate α, β and u.

(2) Determine the optimal value of t:

    (a) if $\alpha \leq t_1$, we can choose $t = \alpha$.

    (b) else, i.e., if $\alpha > t_1$ then $t = t_1$.

(3) Calculate $\lambda_{\alpha,\beta} = \lambda(t; \alpha, \beta, u)$.

(4) Repeat step 2 and 3 interchanging α by β and β by α. The result will be the $\lambda_{\beta,\alpha} = \lambda(t; \beta, \alpha, u)$.

(5) If $\lambda_{\alpha,\beta} = \lambda_{\beta,\alpha}$, go again to step 1. Else, we have our examples.

```python
from math import sqrt, ceil;
#Definition of the formula for the eigenvalues.
#There is an optional parameter, mode,
#which one can change to give us only the
#+-solutions or the maximum of both eigenvalues.
#By default it is set to max.
def eig(alpha,beta,t,u,mode = 'max'):
    #calculate the part with and
    #without the square root of the formula.
    no_sqrt = (alpha - t + 2 - beta)/2;
    yes_sqrt = (sqrt((beta+alpha-t)**2-4*beta*(alpha-t)*u**2))/2;
    #cases considering the different values of the mode parameter.
    if mode == 'plus':
        value = no_sqrt+yes_sqrt;
        return value;
    elif mode == 'minus':
        value = no_sqrt-yes_sqrt;
        return value;
    elif mode == 'max':
        #this will return the norm, i.e. the maximum of the eigenvalues.
        value = max(no_sqrt+yes_sqrt, no_sqrt-yes_sqrt);
        return value;
    else:
        return False;
def t1(c, u):
    if u==0:
        return 1;
```

```
    else:

        return (1-c)/(1-c + c*u**2);
#This is to calculate the constraint,
#we just compare t1(beta) and alpha
def tconstraint(alpha,beta,u):
    t1ab = t1(beta,u);
    t1ba = t1(alpha,u);
    if alpha <=t1ab:
        constraint_ab = alpha;
    else:
        constraint_ab = t1ab;
    if beta <= t1ba:
        constraint_ba = beta;
    else:
        constraint_ba = t1ba;
    #We return the values in the format of a dictionary
    return {'constraint_ab':constraint_ab,
     'constraint_ba':constraint_ba};
#This function will look for suitable points for
# which the norm varies in dependence of
#the permutation of alpha and beta.
#If count is set to true then we will count the
#numbers that are in the different
#sets M<<, M<> or M>>
def look_for_points(N,all = False, count = False):
    #we create a grid in (0,1) of N-2 points.
    #The precision we will use is 5 decimal places
    precision = 5;
    #the number of total points (alpha,beta,u) which are looked
    total_points = (N+1)*(N+1)*N;
    #help variable for the percentage process
    perc_step = 0;
    count_total = 0;
    count_dif =0;
    count_great_great=0;
    count_great_less=0;
    count_less_less=0;
    count_less_great=0;
```

```
eq_count_great_great=0;
eq_count_great_less=0;
eq_count_less_less=0;
eq_count_less_great=0;
for i in range(0,N+1):
    for j in range(0,N+1):
        #u is in [0,1)
        for k in range(1,N+1):
            alpha = round((1-i/N)*1 + (i/N)*0,precision);
            beta = round((1-j/N)*1 + (j/N)*0,precision);
            u =round((1-k/N)*1 + (k/N)*0,precision);
            constraint = tconstraint(alpha,beta,u);
            tab = round(constraint['constraint_ab'],precision);
            tba = round(constraint['constraint_ba'],precision);
            ab = round(eig(alpha, beta, tab,u),precision);
            ba = round(eig(beta,alpha,tba,u),precision);
            count_total+=1;
            #Percentage process
            if count_total/total_points >= perc_step:
                print(str(ceil(100*perc_step))+'%');
                perc_step += .1;
            if ab != ba:
                count_dif+=1;
                points = {'alpha':alpha, 'beta':beta, 'u':u,
                 'tab':tab, 'tba':tba, 'ab':ab, 'ba':ba};
                if all:
                    if alpha > t1(beta,u):
                        if beta > t1(alpha,u):
                            count_great_great+=1;
                        elif beta < t1(alpha,u):
                            count_great_less+=1;
                    elif alpha < t1(beta,u):
                        if beta > t1(alpha,u):
                            count_less_great+=1;
                        elif beta < t1(alpha,u):
                            count_less_less+=1;
                else:
                    #When we only want a suitable point,
```

```
                        #we get the first suitable point.
                        return points;
                else:
                        if all:
                                if alpha > t1(beta,u):
                                        if beta > t1(alpha,u):
                                                eq_count_great_great+=1;
                                        elif beta < t1(alpha,u):
                                                eq_count_great_less+=1;
                                elif alpha < t1(beta,u):
                                        if beta > t1(alpha,u):
                                                eq_count_less_great+=1;
                                        elif beta < t1(alpha,u):
                                                eq_count_less_less+=1;
        if(count and all):
                print('--------------------');
                print('total points = '+str(count_total));
                print('\ntotal suitable points = '+str(count_dif));
                print('   total great great points = '+str(count_great_great));
                print('   total great less points = '+str(count_great_less));
                print('   total less great points = '+str(count_less_great));
                print('   total less less points = '+str(count_less_less));
                print('\ntotal nonsuitable points = '+str(count_total-count_dif));
                print('   total great great points = '+str(eq_count_great_great));
                print('   total great less points = '+str(eq_count_great_less));
                print('   total less great points = '+str(eq_count_less_great));
                print('   total less less points = '+str(eq_count_less_less));
        return False;
look_for_points(10,all=True, count = True)
```

References

[1] Beneduci, Roberto: *Joint measurability through naimark's theorem*. arXiv preprint arXiv:1404.1477, 2014.

[2] Heinosaari, Teiko, Daniel Reitzner, and Peter Stano: *Notes on joint measurability of quantum observables*. Foundations of Physics, 38(12):1133–1147, 2008.

[3] Kunjwal, Ravi, Chris Heunen, and Tobias Fritz: *All joint measurability structures are quantum realizable*. arXiv preprint arXiv:1311.5948, 2013.

[4] Ludwig, Dr. Rer. Nat. Günther: *Die Grundlagen der Quantenmechanik*. Die Grundlehren der Mathematischen Wissenschaften 70. Springer Berlin Heidelberg, 1954, ISBN 978-3-662-11921-1,978-3-662-11920-4.

[5] Ludwig, G: *Der meßprozeß*. Zeitschrift für Physik A Hadrons and Nuclei, 135(5):483–511, 1953.

[6] Lysternik, LA and VI Sobolev: *Elemente der Funktionalanalysis*. Akademie Verlag, 1968.

[7] Vandenberghe, Lieven and Stephen Boyd: *Semidefinite programming*. SIAM review, 38(1):49–95, 1996.

[8] Wolf, Michael M., David Perez-Garcia, and Carlos Fernandez: *Measurements incompatible in quantum theory cannot be measured jointly in any other no-signaling theory*. Phys. Rev. Lett., 103:230402, Dec 2009. http://link.aps.org/doi/10.1103/PhysRevLett.103.230402.