Order preserving maps on quantum measurements

Teiko Heinosaari[1,2], Maria Anastasia Jivulescu[3], and Ion Nechita[4]

[1]Quantum algorithms and software, VTT Technical Research Centre of Finland Ltd

[2]Department of Physics and Astronomy, University of Turku, Finland

[3]Department of Mathematics, Politehnica University of Timişoara, Romania

[4]Laboratoire de Physique Théorique, Université de Toulouse, CNRS, UPS, France

We study the partially ordered set of equivalence classes of quantum measurements endowed with the post-processing partial order. The post-processing order is fundamental as it enables to compare measurements by their intrinsic noise and it gives grounds to define the important concept of quantum incompatibility. Our approach is based on mapping this set into a simpler partially ordered set using an order preserving map and investigating the resulting image. The aim is to ignore unnecessary details while keeping the essential structure, thereby simplifying e.g. detection of incompatibility. One possible choice is the map based on Fisher information introduced by Huangjun Zhu, known to be an order morphism taking values in the cone of positive semidefinite matrices. We explore the properties of that construction and improve Zhu's incompatibility criterion by adding a constraint depending on the number of measurement outcomes. We generalize this type of construction to other ordered vector spaces and we show that this map is optimal among all quadratic maps.

Contents

Teiko Heinosaari: teiko.heinosaari@utu.fi

Maria Anastasia Jivulescu: maria.jivulescu@upt.ro

Ion Nechita: ion.nechita@univ-tlse3.fr

1   Introduction

There is a general guideline in mathematics how to study a complicated structure: map it to something simpler while keeping the essential structural features of it. Then, by investigating the simple structure one can get information also on the more complicated structure. In the present work, we follow this guideline to investigate the post-processing order of quantum measurements. Post-processing can be defined as a classical manipulation of measurement outcome data, but it can also be arising e.g. from weak measurement coupling, disturbance caused by environment or some other source of noise. Therefore, post-processing ordering does not mean that there has been active processing of measurement data, but rather that in principle two measurements could be related in that specific way. Generally, one can say that the post-processing order describes which measurement is noisier than another measurement. An important aspect is that the post-processing order provides an operational definition of quantum incompatibility, which has become central in several quantum information processing applications. The post-processing relation was initially studied in [Md90] and further results and developments have been reported e.g. in [Hei05, BDK$^+$05, JP07, JPV08, ACHT09, HHP12, Kur15a, Kur15b, HP17, GMSG21].

The starting point of our investigation is that we map the set of all quantum measurements to a partially ordered vector space by an order preserving map. We demonstrate how quadratic maps on matrices lead to this kind of setting in a natural way. This approach has been used by Huangjun Zhu [Zhu15, ZHC16], who defined a particular kind of map related to Fisher information. We show that the map introduced by Zhu is exceptional among all quadratic maps and we explore the properties of that specific construction. We also give examples of other order preserving maps and show that some of them can be useful when investigating the post-processing order of quantum measurements, although they cannot detect incompatibility. The simpler the map is, the less information it gives. But a complicated map does not simplify the structure enough and hence does not provide any actual relief. From this respect, Zhu's map is in good balance and that explains its special status and usefulness for concrete calculations.

The paper is organized as follows. In Section 2 we explain the core idea in a general setting of preordered sets. In the rest of the paper we then limit to quantum measurements. Their post-processing order is recalled Section 3. Then, in Section 4 we explain a method to construct order preserving maps on the set of quantum measurements and we provide several examples. After presenting the general framework, in Section 5 we go into the details of Zhu's construction. In Section 5.3 we recall and improve Zhu's incompatibility criterion by adding a constraint depending on the number of outcomes of the quantum measurement. Finally, in Section 6 we provide several examples that illustrate the improved
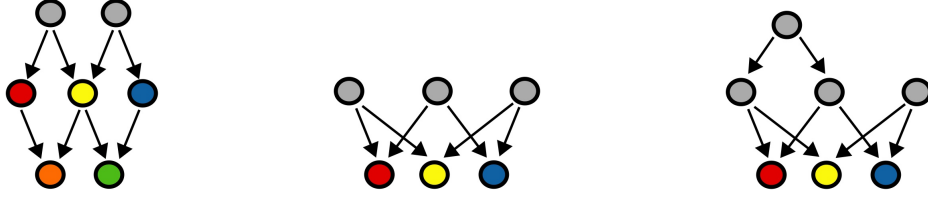
Figure 1: From left to right: (a) All six colored elements are pairwise compatible except the pair consisting red and blue elements. (b) All colored elements are pairwise compatible but the colored triplet is incompatible. (c) The colored triplet is compatible.

Zhu criterion for incompatibility.

2 Prelude: (in)compatibility in preordered sets

To get a structural perspective to post-processing and incompatibility, we start with a general setting of preordered sets.

We recall that a *preorder* on a set Ω is a binary relation $\leq$ that is reflexive and transitive. For a nonempty subset $X \subseteq \Omega$, we denote

$$\uparrow X = \{ \, y \in \Omega : x \leq y \quad \forall x \in X \, \},$$
$$\downarrow X = \{ \, y \in \Omega : y \leq x \quad \forall x \in X \, \}.$$

A subset X is called *compatible* if $\uparrow X \neq \emptyset$, i.e., there exists an element $y \in \Omega$ such that $x \leq y$ for all $x \in X$; otherwise X is called *incompatible*. We also say that the elements in X are (in)compatible when X is (in)compatible. A preorderer set can be visualized as a directed graph, with elements of the set corresponding to vertices while the preorder between two elements corresponds to the directed edges between vertices. Illustrations are given in Fig. 1.

The operational interpretation of compatibility and incompatibility depends on the preorder. In this section, we make some general definitions and we further illustrate the concept with some simple examples. Our primary interest, namely, the post-processing preorder of quantum measurements, is defined later in Sec. 3.

We recall the following notions. An element $x \in \Omega$ is

- *maximal* if $x \leq y$ implies that $y \leq x$ for all $y \in \Omega$.

- *minimal* if $y \leq x$ implies that $x \leq y$ for all $y \in \Omega$.

There can be many maximal and minimal elements (see Fig. 1), although in exceptional cases they are unique and then called the greatest and the least element, respectively. We note that two maximal elements x and y are compatible if and only if $x \leq y \leq x$.

Example 2.1. Let Ω be a convex set. For two elements $x, y \in \Omega$, we define $x \leq y$ if y is in some nontrivial convex decomposition of x, i.e., $x = ty + (1 - t)y'$ for some $0 < t < 1$ and $y' \in \Omega$. The maximal elements are, by definition, the extreme elements of Ω. Any collection of different extreme elements is incompatible.

We recall that a preorder $\leq$ is a *partial order* if it is antisymmetric, i.e., $x \leq y \leq x$ implies $x = y$. The usual operator ordering, recalled in the next example, is a partial order and we will use it later.

Example 2.2. Let $\mathcal{H}$ be a complex Hilbert space and $\mathcal{L}_+(\mathcal{H})$ the set of all positive operators on $\mathcal{H}$, where an operator T is *positive* if $\langle \psi \,|\, T\psi \rangle \geq 0$ for all $\psi \in \mathcal{H}$. Positive operators correspond to positive semidefinite matrices. The usual operator ordering is defined as $S \leq T$ if $T - S$ is positive, and it is a partial order. For any $S, T \in \mathcal{L}_+(\mathcal{H})$, we have $S \leq S + T$ and $T \leq S + T$, hence $S + T$ is an upper bound for S and T. However, it is known that $\mathcal{L}_+(\mathcal{H})$ is not a lattice but an antilattice; the least upper bound for S and T exists if and only if $S \leq T$ or $T \leq S$ [Kad51].

For our purposes, there is no big difference in preorder and partial order. Namely, a preorder $\leq$ determines an equivalence relation $\simeq$ on Ω by setting $x \simeq y$ if $x \leq y \leq x$. We denote by $[x]$ the equivalence class of x. Further, $\leq$ determines a partial order on the set of all equivalence classes as follows: $[x] \leq [y]$ if and only if $x \leq y$ (without causing confusion we denote the induced partial order with the same symbol $\leq$). We denote by $\Omega^\simeq$ this partially ordered set (poset).

Importantly, it follows from the definition of compatibility that $X \subseteq \Omega$ is compatible if and only if $X^{[]} \equiv \{[x] : x \in X\}$ is compatible. We conclude that there is no essential difference if we investigate compatibility in a preordered set or in the induced partially ordered set.

Example 2.3. Let $\mathcal{P}(n)$ be the convex set of probability distributions on the set $\{1, \ldots, n\}$. We consider the preorder defined in Example 2.1. We can interpret the Kronecker distributions δ_j as describing precise knowledge, whereas other probability distributions describe knowledge with some uncertainty. For $p, q \in \mathcal{P}(n)$, we have $p \leq q$ if and only if $\mathrm{supp}\,(p) \supseteq \mathrm{supp}\,(q)$. We see that $p, q \in \mathcal{P}(n)$ are compatible if and only if there exists $j \in \{1, \ldots, n\}$ such that $p \leq \delta_j$ and $q \leq \delta_j$, which is the case when $p(j)q(j) \neq 0$. Hence, compatibility of p and q means that they may origin from the same precise description δ_j, just the uncertainty is different. We have $p \simeq q$ if and only if $\mathrm{supp}\,(p) = \mathrm{supp}\,(q)$. The maximal elements in $\mathcal{P}(n)^\simeq$ are the singleton sets $\{\delta_j\}$, $j = 1, \ldots, n$, and the least element is the subset containing all probability distributions with full support.

The main tool in our later developments will be order preserving maps between two preordered (or partially ordered) sets. Let $(\Omega, \leq)$ and $(\Omega', \leq')$ be two preordered sets. A map $h : \Omega \to \Omega'$ is order preserving if

$$x \leq y \quad \Rightarrow \quad h(x) \leq' h(y)$$

for all $x, y \in \Omega$. We make the following simple observations.

Proposition 2.4. *Let h be an order preserving map. For any $X \subseteq \Omega$, the following implications hold:*

(a) *If X is compatible, then the image $h(X)$ is compatible.*

(b) *If the image $h(X)$ is incompatible, then X is incompatible.*

The latter observation is the starting point for our investigation and it is illustrated in Fig. 2(a). This observation means that we can obtain sufficient criteria for incompatibility on Ω by investigating other preordered sets and relevant order preserving maps.

Suppose we want to investigate a preordered set $(\Omega, \leq)$ and for that aim we choose sets $(\Omega', \leq')$ and order preserving maps $h : \Omega \to \Omega'$. There is an obvious trade-off between lost information and difficulty. Namely, simpler is the range $h(\Omega)$, easier it is to decide if the image $h(X)$ of a subset X is incompatible. In contrast, in a simpler case we loose more about the original structure of $(\Omega, \leq)$ and hence detect less incompatible sets.
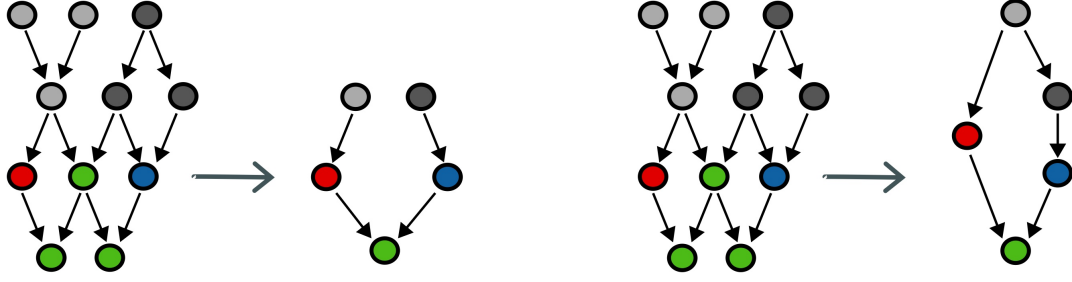
Figure 2: A partially order set is mapped to another partially order set by an order preserving map. Similarly colored elements are mapped into respectively colored elements. From left to right: (a) In the image side the incompatibility of the red and blue elements is evident, hence they are incompatible also on the domain side. (b) The image contains the greatest element. Therefore, all the elements on the image side are compatible and the map is useless for incompatibility detection.

Example 2.5. Let $(\Omega, \leq)$ and $(\Omega', \leq')$ be two preordered sets and $h : \Omega \to \Omega'$ an order preserving map. Suppose that $h(\Omega)$ has the greatest element. It follows that $h(X)$ is compatible for any $X \subset \Omega$ and Prop. 2.4(b) is therefore useless in detecting incompatibility. This kind of situation is depicted in Fig. 2(b).

3 Post-processing order of quantum measurements

We review in this section the basic theory of quantum measurements, using the formalism of Positive Operator Valued Measures (POVMs), also known as quantum meters. Our focus is on the partial ordering induced by *classical post-processing*, and the related, crucial phenomenon of *quantum incompatibility*.

3.1 Quantum measurements

In the following $\mathcal{H}$ is a fixed finite d-dimensional Hilbert space. We denote by $\mathcal{S}(\mathcal{H})$ the set of all states, i.e., positive trace class operators of trace 1 and by $\mathcal{E}(\mathcal{H})$ the set of effects, that are linear operators E on $\mathcal{H}$ such that $0 \leq E \leq \mathbb{1}$. We will restrict our investigation to quantum measurements with a finite number of outcomes, hence we will understand a measurement as a map $\mathsf{A} : x \to \mathsf{A}_x$ from a finite set of measurement outcomes Ω_{A} to the set of bounded linear operators $\mathcal{L}(\mathcal{H})$ on $\mathcal{H}$ such that $\mathsf{A}_x \geq 0$ and $\sum_x \mathsf{A}_x = \mathbb{1}$. This is also known as a positive operator valued measure (POVM). For a subset $X \subseteq \Omega_{\mathsf{A}}$, we denote $\mathsf{A}_X = \sum_{x \in X} \mathsf{A}_x$. The probability of getting an outcome x in a measurement of A in an initial state ϱ is given by the formula $\mathrm{tr}\,[\varrho \mathsf{A}_x]$. We denote by $\mathcal{O}_d$ the set of all measurements A on $\mathcal{H}$ with a finite outcome set $\Omega_{\mathsf{A}} \subset \mathbb{Z}$.

3.2 Post-processing preorder

We recall in this section the post-processing order relation on the set of quantum measurements (or POVMs). A matrix $(\mu_{xy})_{x \in X, y \in Y}$ is called *column stochastic* if it has non-negative elements and its column sums are 1:

$$\forall y \in Y, \qquad \sum_{x \in X} \mu_{xy} = 1.$$

Definition 3.1. For two measurements A and B, we denote $A \preccurlyeq B$ if there exists a column stochastic matrix μ such that

$$\forall x \in X, \qquad A_x = \sum_{y \in Y} \mu_{xy} B_y. \tag{1}$$

We call this relation *post-processing*.

It is straightforward to verify that $\preccurlyeq$ is a preorder.

Example 3.2. A special type of post-processing is *relabeling*, which means that $\mu_{xy} \in \{0, 1\}$ for all x, y. In this case, μ being column stochastic implies that there is a relabeling function $r : X \to Y$ such that (1) can be written as

$$\forall x \in X, \qquad A_x = \sum_{y \in r^{-1}(x)} B_y. \tag{2}$$

In this special case, we say that A is a relabeling of B and that B is a refinement of A.

In the following, we summarize some known results that are relevant for our investigation. The proofs can be found in [Md90].

Proposition 3.3. *(a) A measurement $A \in \mathcal{O}_d$ is post-processing minimal if and only if A is trivial, i.e. each $A_x = a_x \mathbb{1}$ for some $a_x \in \mathbb{R}$.*

(b) A measurement $A \in \mathcal{O}_d$ is post-processing maximal if and only if A is rank-1, i.e. each nonzero A_x is a rank-1 operator.

3.3 Equivalence classes and simple representatives

The post-processing preorder relation introduced above is not anti-symmetric, meaning that there exist distinct measurements $A \neq B$ such that both $A \preccurlyeq B$ and $B \preccurlyeq A$ hold. To turn $\preccurlyeq$ into a partial order, we need to quotient it by the corresponding equivalence classes as explained in Section 2.

Definition 3.4. Given two measurements A and B, we denote $A \sim B$ if $A \preccurlyeq B \preccurlyeq A$, and in this case we say that A and B are *post-processing equivalent*. We denote by $[A]$ the equivalence class of an measurement A, and by $\mathcal{O}_d^{\sim}$ the set of all equivalence classes of measurements on $\mathcal{H} \cong \mathbb{C}^d$.

The relation $\preccurlyeq$ induces a *partial order* (which we also denote by $\preccurlyeq$) on the set $\mathcal{O}_d^{\sim}$ of equivalence classes of measurements.

It is convenient to be able to choose a "standard representative" of an equivalence class $[A]$ of measurements when needed. We say that a measurement A is *simple* if every $B \in [A]$ is a refinement of A. It has been shown in [Kur15b] that every equivalence class has a simple representative (called minimally sufficient in [Kur15b]) that can be formed iteratively as follows: start from A and merge two operators A_x and A_y if there exists $c \geq 0$ such that $A_x = cA_y$. Then continue this merging until there are no more such pairs. It follows that a measurement A is simple if and only if, whenever $x, y \in \Omega_A$ and $x \neq y$, there is no $c \geq 0$ such that $A_x = cA_y$. Further, it can be shown [Kur15b] that if $A \sim B$ are both simple, then they have the same number of outcomes and there is bijective relabeling that connects them. We denote by $\ell([A])$ the number of POVM elements of the simple representative of $[A]$.
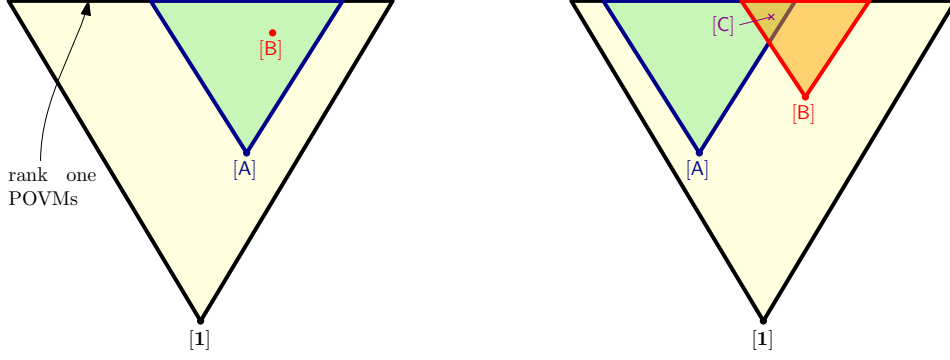
Figure 3: The poset of equivalence classes of quantum measurements has the least element $[\mathbb{1}]$ and infinitely many maximal elements. Left: $[\mathsf{A}] \prec [\mathsf{B}]$. Right: compatible measurements A and B with a joint measurement C.

Remark 3.5. We conclude from the previous discussion that the equivalence classes are in one-to-one correspond to finite subsets of effects $\{E_1, E_2, \ldots, E_n\} \subset \mathcal{E}(\mathcal{H})$ such that

(i) $\{E_1, E_2, \ldots, E_n\}$ does not contain collinear operators;

(ii) $\sum_i E_i = \mathbb{1}$.

Hence, we can identify $\mathcal{O}_d^{\sim}$ with the collection of these type of subsets of $\mathcal{E}(\mathcal{H})$ whenever it is convenient.

The structure of the partially ordered set $(\mathcal{O}_d^{\sim}, \preccurlyeq)$ is governed by the fact that the equivalence class of trivial POVMs(denoted by $[\mathbb{1}]$) is the least element and the maximal elements are those $[\mathsf{A}]$ with its simple representative A having $\mathrm{rank}(\mathsf{A}_i) = 1$ for all $i \in \Omega_\mathsf{A}$ (some other representative consists also of rank-1 operators apart from zero effects); see Figure 3, the left panel. We further endow the set $\mathcal{O}_d^{\sim}$ with a convex structure as follows: for two equivalence classes $[\mathsf{A}]$ and $[\mathsf{B}]$ and $\lambda \in [0, 1]$, we define

$$\lambda[\mathsf{A}] + (1 - \lambda)[\mathsf{B}] := [\lambda \mathsf{A} \sqcup (1 - \lambda)\mathsf{B}]. \tag{3}$$

In terms of the simple representatives, the convex combination of two equivalence classes is given by the class of the measurement obtained by taking the (weighted) *union* of the effects from each class. We remark that $\lambda \mathsf{A} \sqcup (1 - \lambda)\mathsf{B}$ is a different notion of convex mixture than the usual one for measurements, where one mixes corresponding effects of two measurements with equal number of outcomes. The definition above does not depend on the representatives chosen: assuming that $\mathsf{A}' \sim \mathsf{A}$ and $\mathsf{B}' \sim \mathsf{B}$, we observe that $\lambda \mathsf{A} \sqcup (1 - \lambda)\mathsf{B} \sim \lambda \mathsf{A}' \sqcup (1 - \lambda)\mathsf{B}'$ by taking direct sums of the respective column stochastic matrices.

3.4 (In)compatibility of measurements

As defined generally in Section 2, a set of measurements $\mathcal{A} \subseteq \mathcal{O}_d$ is compatible if there exists a measurement C such that $\mathsf{A} \preccurlyeq \mathsf{C}$ for every $\mathsf{A} \in \mathcal{A}$; otherwise $\mathcal{A}$ is incompatible. In this context, compatibility has an interpretation as a joint measurement, also called simultaneous measurement. Equivalently (see [ACHT09]), two measurements A and B are compatible if it exists a third measurement C, with outcome set $X \times Y$, such that $\sum_y \mathsf{C}_{xy} = \mathsf{A}_x, \forall x$ and $\sum_x \mathsf{C}_{xy} = \mathsf{B}_y, \forall y$. In this case, post-processing from C to A and B amounts to calculating the corresponding marginals; see Figure 4.

We can also introduce compatibility for equivalence classes of quantum measurements.
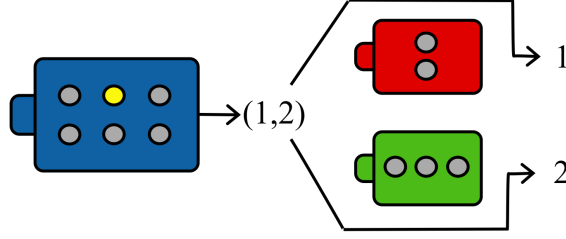
Figure 4: In a joint measurement, one measurement device (blue) simulates two given measurement devices (red and green) by giving their outcomes as marginals. In this particular case the blue device gives the outcome $(1, 2)$, from which one concludes that the outcome of the red device is 1 and that of the green device is 2.

Definition 3.6. A set $\mathcal{A}^{[]} \subseteq \mathcal{O}_d^{\sim}$ of equivalence classes of quantum measurements is *compatible* if there exists a class $[\mathsf{C}] \in \mathcal{O}_d^{\sim}$ such that $[\mathsf{A}] \preccurlyeq [\mathsf{C}]$ for all $[\mathsf{A}] \in \mathcal{A}^{[]}$.

This definition falls again into the general setting discussed in Section 2: two equivalence classes of measurements $[\mathsf{A}]$ and $[\mathsf{B}]$ are compatible if there exists an element $[\mathsf{C}]$ in $(\mathcal{O}_d^{\sim}, \preccurlyeq)$ larger than both of them; see Figure 3, right panel. As compatibility is defined via post-processing, it is trivial that $[\mathsf{A}]$ and $[\mathsf{B}]$ are compatible iff A and B are compatible.

Example 3.7. Generally, commutativity is a sufficient condition for compatibility but not necessary. It is a well-known fact that a quantum measurement A that consists of projections (i.e. $\mathsf{A}_x^2 = \mathsf{A}_x \ \forall x$) is compatible with another measurement B if and only if they commute, $\mathsf{A}_x \mathsf{B}_y = \mathsf{B}_y \mathsf{A}_x \ \forall x, y$ (see e.g. [HRS08]). We conclude that the same is true for any measurement $\mathsf{A}' \in [\mathsf{A}]$. For instance, A' can have effects $\mathsf{A}_1' = a_1 \mathsf{A}_1$, $\mathsf{A}_2' = a_2 \mathsf{A}_1$ with $a_1 + a_2 = 1$ and $0 < a_1 < 1$. In this case A_1' is a not a projection but $\mathsf{A}' \in [\mathsf{A}]$ and the statement about the equivalence between commutativity and compatibility holds therefore for A'.

4 Construction of order preserving maps

In this section, we introduce the notion of *order preserving maps*, the main tool we shall use to study the post-processing partial order and the compatibility structure of quantum measurements. The main idea behind these considerations is that we want to map the partially order set of quantum measurements to a different poset, which has, simultaneously, the following two desirable properties:

- it is simpler, allowing for practical criteria for non-ordering and incompatibility

- it still captures enough of the structural properties of the quantum measurement poset to allow for powerful non-ordering and incompatibility criteria.

We will present several maps and make it precise how they can be compared.

4.1 General construction

As mentioned already in Section 2, we will study the partially ordered set $(\mathcal{O}_d^{\sim}, \preccurlyeq)$ by studying morphisms to simpler posets, such as the set of positive semidefinite matrices. In the following, we describe a general method to define order preserving maps from $(\mathcal{O}_d^{\sim}, \preccurlyeq)$ into a partially ordered vector space. Let $(\mathcal{V}, \leq)$ be a partially ordered vector space, i.e., $\mathcal{V}$ is a real vector space, and $\leq$ is a relation on $\mathcal{V}$ satisfying the following two conditions

- $x \leq y \implies x + z \leq y + z$ for all $x, y, z \in \mathcal{V}$

- $x \leq y \implies \lambda x \leq \lambda y$ for all $x, y \in \mathcal{V}$ and $\lambda \in \mathbb{R}, \lambda \geq 0$.

The set $\mathcal{L}_s(\mathcal{H})$ of selfadjoint operators with the usual operator ordering is a paradigmatic example of a partially ordered vector space. The following notion of order morphism is crucial for our investigation.

Definition 4.1. A map $g : \mathcal{L}_s(\mathcal{H}) \to \mathcal{V}$ such that

(i) $g(\lambda E) = \lambda g(E)$ for all scalars $\lambda \geq 0$.

(ii) $g(E + F) \leq g(E) + g(F)$

is called an *order morphism*. To any order morphism, we associate a map $G : \mathcal{O}_\mathcal{H} \to \mathcal{V}$ given by

$$G(\mathsf{A}) = \sum_x g(\mathsf{A}_x) \,. \tag{4}$$

Remark 4.2. The property (ii) above corresponds to the definition of *operator convex functions*, see [Bha97, Section 5]. However, in that setting, the function $g(E)$ is defined using functional calculus (i.e. g is a real function which acts on the eigenvalues of the input), whereas we allow for more general matrix functions. In particular, asking that the property (i) above holds for a real function implies that the function is linear.

We note that in order for the map G to be interesting and capture some of the order-related properties of the set $\mathcal{O}_\mathcal{H}$, the map g must be *non-linear*. Indeed, if the map g is linear, then we have

$$G(\mathsf{A}) = \sum_x g(\mathsf{A}_x) = g(\sum_x \mathsf{A}_x) = g(\mathbb{1})$$

and therefore G is constant on $\mathcal{O}_\mathcal{H}$.

The following is central for our investigation and links G to our general discussion in Section 2 .

Proposition 4.3. *The map G defined in (4) is order preserving on $\mathcal{O}_d$, i.e., $\mathsf{A} \preccurlyeq \mathsf{B}$ implies that $G(\mathsf{A}) \leq G(\mathsf{B})$ for all $\mathsf{A}, \mathsf{B} \in \mathcal{O}_d$. Therefore, G respects the equivalence relation of quantum measurements and it induces a map $G : \mathcal{O}_d^\sim \to \mathcal{V}$. The induced map G is order preserving on $\mathcal{O}_d^\sim$.*

Proof. Let us first show the statement for $(\mathcal{O}_d, \preccurlyeq)$. Given two measurements A, B with $\mathsf{A} \preccurlyeq \mathsf{B}$, there exists a column stochastic matrix μ such that (1) holds. We then have

$$G(\mathsf{A}) = \sum_x g(\mathsf{A}_x) = \sum_x g\left(\sum_y \mu_{xy} \mathsf{B}_y\right) \leq \sum_x \sum_y \mu_{xy} g(\mathsf{B}_y) = \sum_y \underbrace{\left(\sum_x \mu_{xy}\right)}_{=1} g(\mathsf{B}_y) = \sum_y g(\mathsf{B}_y) = G(\mathsf{B}).$$

It follows that $\mathsf{A} \preccurlyeq \mathsf{B} \preccurlyeq \mathsf{A}$ implies $G(\mathsf{A}) = G(\mathsf{B})$, hence G respects the equivalence relation. The last statement is a consequence of these two facts. $\qquad\square$

Proposition 2.4(b) can be written in the following form in the present setting.

Proposition 4.4. *Let $\mathcal{A}^{\mathbb{1}} \subset \mathcal{O}_d^\sim$. If there is no element $H \in G(\mathcal{O}_d^\sim)$ such that $G([\mathsf{A}]) \leq H$ for all $[\mathsf{A}] \in \mathcal{A}^{\mathbb{1}}$, then $\mathcal{A}$ is incompatible.*

Deciding incompatibility of a collection of measurements by using Prop. 4.4 requires that we know the range of G, or at least something about it. In this respect, an important fact is that he map G is affine with respect to the convex structure of $\mathcal{O}_d^\sim$ introduced in (3).

Proposition 4.5. *Given two measurements* A *and* B*, as well as a convex weight* $\lambda \in [0, 1]$*, the map* G *respects the convex structure of the triangle introduced in* (3)*, i.e.,*

$$G(\lambda[\mathsf{A}] + (1-\lambda)[\mathsf{B}]) = \lambda G([\mathsf{A}]) + (1-\lambda)G([\mathsf{B}]).$$

In particular, the range of the map G *is a convex set.*

Proof. The fact that G is affine can be checked by using representatives of the classes $[\mathsf{A}], [\mathsf{B}]$:

$$G(\lambda[\mathsf{A}] + (1-\lambda)[\mathsf{B}]) = \sum_{x=1}^{m} g(\lambda \mathsf{A}_x) + \sum_{y=1}^{n} g((1-\lambda)\mathsf{B}_y) = \lambda G([\mathsf{A}]) + (1-\lambda)G([\mathsf{B}]),$$

where we have used the fact that $\{\lambda \mathsf{A}_x\}_{x=1}^{m} \sqcup \{(1-\lambda)\mathsf{B}_y\}_{y=1}^{n}$ is a representative of the class $\lambda[\mathsf{A}] + (1-\lambda)[\mathsf{B}]$ defined in (3). $\qquad\square$

4.2 Quadratic maps

As noted before, linear maps g are trivial. We focus now on a special class of order preserving maps, namely, that of *quadratic maps*. These are maps g of the form

$$g(E) = \frac{\Phi(E)\Phi(E)^*}{\tau(E)}, \tag{5}$$

where $\Phi : M_d(\mathbb{C}) \to M_{D \times r}(\mathbb{C})$ is a linear map and $\tau : M_d(\mathbb{C}) \to \mathbb{C}$ is a faithful positive linear form; in particular, we have $\tau(E) = \operatorname{tr}[\rho E]$ for some positive definite matrix ρ. (We can relax the last assumption on ρ by requiring that the kernel of τ should be contained in the kernel of Φ.) First, let us show that such maps verify the two axioms for order morphisms.

Proposition 4.6. *The quadratic maps from Eq.* (5) *satisfy the order morphism axioms of Definition 4.1.*

Proof. First, for a positive scalar $\lambda > 0$, we have

$$g(\lambda E) = \frac{\Phi(\lambda E)\Phi(\lambda E)^*}{\tau(\lambda E)} = \frac{\lambda^2 \Phi(E)\Phi(E)^*}{\lambda \tau(E)} = \lambda g(E)$$

and hence the first axiom is satisfied. Second, we have

$$
\begin{aligned}
\tau(E + F) \, & (g(E) + g(F) - g(E + F)) \\
&= \frac{\tau(F)}{\tau(E)} \Phi(E)\Phi(E)^* + \frac{\tau(E)}{\tau(F)} \Phi(F)\Phi(F)^* - \Phi(E)\Phi(F)^* - \Phi(F)\Phi(E)^* \\
&= \left[\sqrt{\frac{\tau(F)}{\tau(E)}} \Phi(E) - \sqrt{\frac{\tau(E)}{\tau(F)}} \Phi(F) \right] \left[\sqrt{\frac{\tau(F)}{\tau(E)}} \Phi(E) - \sqrt{\frac{\tau(E)}{\tau(F)}} \Phi(F) \right]^* \geq 0 \,,
\end{aligned}
$$

showing that the second axiom is satisfied and concluding the proof. $\qquad\square$

A standard choice for the form τ appearing in the denominator in (5) is $\tau(E) = \mathrm{tr}\,[E]$. Regarding the choice of the linear map Φ in the numerator, several interesting choices can be made:

(LIN1) $\Phi(E) = |E\rangle \in \mathbb{C}^{d^2}$

(LIN2) $\Phi(E) = E$

(LIN3) $\Phi(E) = \mathrm{diag}(E)$

(LIN4) $\Phi(E) = E\psi$ for a fixed unit vector $\psi \in \mathcal{H}$

(LIN5) $\Phi(E) = \mathrm{tr}\,[E]$.

Here $|E\rangle$ denotes the vectorization of a matrix E, that is, the vector obtained by stacking the columns of E on top of each other. More exactly,

$$|E\rangle = \sum_{j,k=1}^{d} \langle k|E|j\rangle |j\rangle \otimes |k\rangle \in \mathbb{C}^{d^2},$$

where $\{|j\rangle\}_{j=1}^{d}$ is an orthonormal basis on $\mathbb{C}^d$. This map is an incarnation of the standard isomorphism between the vector space of linear maps between two vector spaces and their tensor product [BŻ06, Section 10.2]. The first choice above corresponds to the *Fisher information map* introduced by Huangjun Zhu [Zhu15, ZHC16] and will denote the respective quadratic map as f and the derived order preserving map as F. This map will be discussed in detail in Section 5. Let us discuss, in reverse order, the other four examples above. In the following we use $\tau(E) = \mathrm{tr}\,[E]$ in the denominator if not otherwise stated.

- The map $\Phi(E) = \mathrm{tr}\,[E]$ is not a constant map, but it still gives a trivial map G on $\mathcal{O}_d$ as $G(\mathsf{A}) = d$ for any measurement A. It is hence useless for any purposes to study the order structure of $\mathcal{O}_{\tilde{d}}$ and we listed it to illustrate the fact that we cannot reduce the structure too much.

- The choice $\Phi(E) = E\psi$, where $\psi \in \mathcal{H}$ is a fixed unit vector, leads to the map

$$G_\psi(\mathsf{A}) = \sum_x \frac{\mathsf{A}_x|\psi\rangle\langle\psi|\mathsf{A}_x^*}{\langle\psi|\mathsf{A}_x|\psi\rangle}, \tag{6}$$

where we have used the adapted choice $\tau(E) = \langle\psi|\mathsf{A}_x|\psi\rangle$. For the minimal class $[\mathbb{1}]$, we have $G_\psi([\mathbb{1}]) = |\psi\rangle\langle\psi|$, while for a rank-1 measurement $\mathsf{A} = (\mathsf{A}_x = \lambda_x|a_x\rangle\langle a_x|)_x$, we have

$$G_\psi(\mathsf{A}) = \sum_x \lambda_x \frac{|a_x\rangle\langle a_x| \cdot |\psi\rangle\langle\psi| \cdot |a_x\rangle\langle a_x|}{|\langle a_x, \psi\rangle|^2} = \sum_x \lambda_x |a_x\rangle\langle a_x| = \mathbb{1}.$$

Hence, for an arbitrary measurement A, we have $G_\psi(\mathsf{A}) \leq \mathbb{1}$. We conclude that this map cannot detect incompatibility. Namely, $G_\psi(\mathcal{O}_d)$ has the greatest element $\mathbb{1}$ and we are therefore in the situation of Example 2.5.

- For the choice $\Phi(E) = \mathrm{diag}(E)$ the map G reads

$$G(\mathsf{A}) = \sum_x \frac{\mathrm{diag}(\mathsf{A}_x)^2}{\mathrm{tr}\,[\mathsf{A}_x]}. \tag{7}$$

Since, for diagonal matrices, the operator order reduces to coordinate-wise dominance, we obtain a necessary *scalar* condition for the order, namely,

$$[\mathsf{A}] \preccurlyeq [\mathsf{B}] \quad \Rightarrow \quad \forall k \in [d], \quad \sum_x \frac{\mathsf{A}_x(k,k)^2}{\sum_j \mathsf{A}_x(j,j)} \leq \sum_x \frac{\mathsf{B}_x(k,k)^2}{\sum_j \mathsf{B}_x(j,j)}. \tag{8}$$

As in the previous case, $G(\mathcal{O}_d)$ has the greatest element $\mathbb{1}$ hence the map G cannot witness incompatibility of quantum measurements. However, the map G can detect non-order in some particular cases.

- The choice $\Phi(E) = E$ gives $g(E) = E^2/\operatorname{tr}[E]$. This yields the map

$$G(\mathsf{A}) = \sum_x \frac{\mathsf{A}_x^2}{\operatorname{tr}[\mathsf{A}_x]}. \tag{9}$$

This map is easy to implement and, as we will shortly see, it is useful in investigating the order of $\mathcal{O}_d^{\sim}$. On the downside, it cannot detect incompatibility. Namely, for any rank-1 measurement A we have $G(\mathsf{A}) = \mathbb{1}$. Since rank-1 measurements are maximal in $\mathcal{O}_d$, we have $G(\mathsf{B}) \leq \mathbb{1}$ for all $\mathsf{B} \in \mathcal{O}$. It follows that the image of $\mathcal{O}_d^{\sim}$ through this order morphism looks like a rhombus, with $\mathbb{1}$ on top (corresponding to rank-1 measurement) and $\mathbb{1}/d$ on the bottom (corresponding to trivial measurements). This kind of image is useless in incompatibility detection as explained in Example 2.5.

All of the introduced maps provide necessary conditions for two measurements being ordered. Namely, from Prop. 4.3 we conclude that if the matrix $G([\mathsf{A}]) - G([\mathsf{B}])$ has both positive and negative eigenvalues, then neither $[\mathsf{A}] \preccurlyeq [\mathsf{B}]$ nor $[\mathsf{B}] \preccurlyeq [\mathsf{A}]$. This implication is only in one direction, but the test is easy to implement and therefore useful. For instance, let us consider the following two outcome, three outcome and four outcome qutrit measurements:

$$\mathsf{A}_1 = \frac{1}{6} \begin{pmatrix} 5 & 0 & 1 \\ 0 & 4 & -1 \\ 1 & -1 & 3 \end{pmatrix}, \quad \mathsf{A}_2 = \frac{1}{6} \begin{pmatrix} 1 & 0 & -1 \\ 0 & 2 & 1 \\ -1 & 1 & 3 \end{pmatrix}$$

and

$$\mathsf{B}_1 = \frac{1}{12} \begin{pmatrix} 2 & 0 & 1 \\ 0 & 1 & -1 \\ 1 & -1 & 3 \end{pmatrix}, \quad \mathsf{B}_2 = \frac{1}{12} \begin{pmatrix} 4 & -2 & 1 \\ -2 & 7 & 1 \\ 1 & 1 & 7 \end{pmatrix}, \quad \mathsf{B}_3 = \frac{1}{6} \begin{pmatrix} 3 & 1 & -1 \\ 1 & 2 & 0 \\ -1 & 0 & 1 \end{pmatrix}$$

and

$$\mathsf{C}_1 = \frac{1}{12} \begin{pmatrix} 7 & 0 & 2 \\ 0 & 5 & -2 \\ 2 & -2 & 6 \end{pmatrix}, \quad \mathsf{C}_2 = \frac{1}{12} \begin{pmatrix} 1 & -1 & 0 \\ -1 & 3 & 1 \\ 0 & 1 & 2 \end{pmatrix},$$

$$\mathsf{C}_3 = \frac{1}{24} \begin{pmatrix} 5 & 2 & -1 \\ 2 & 4 & 0 \\ -1 & 0 & 1 \end{pmatrix}, \quad \mathsf{C}_4 = \frac{1}{24} \begin{pmatrix} 3 & 0 & -3 \\ 0 & 4 & 2 \\ -3 & 2 & 7 \end{pmatrix}.$$

Application of the map (7) shows that both pairs $[\mathsf{A}], [\mathsf{C}]$ and $[\mathsf{B}], [\mathsf{C}]$ are not ordered, but fails to indicate anything else. The map (9), on the other hand, shows that $[\mathsf{A}], [\mathsf{B}]$ and $[\mathsf{B}], [\mathsf{C}]$ are not ordered, but does not reveal anything else. Finally, using the Fisher information map F shows that all three equivalence classes $[\mathsf{A}]$, $[\mathsf{B}]$ and $[\mathsf{C}]$ are pairwise not comparable.

4.3 Comparing order preserving maps

The difference between the quadratic maps that was noted in the end of last subsection motivates to compare order preserving maps in their power to detect non-orderings. Clearly, if $[A] \preccurlyeq [B]$, then $G([A]) \leq G([B])$ for any order preserving map G. But, as shown in the previous subsection, there are also pairs $[A], [B]$, that are not comparable but still satisfy $G([A]) \leq G([B])$. These are unwanted cases as then G cannot observe the non-ordering. The fewer of these cases we have, the more informative is the order preserving map - this is the meaning of the following definition.

Definition 4.7. Let $G : \mathcal{O}_d^\sim \to \mathcal{V}$ and $G' : \mathcal{O}_d^\sim \to \mathcal{V}'$ be two order preserving maps. We say that G is at least as informative as G' if for all quantum measurements A, B, the implication

$$G([A]) \leq G([B]) \quad \Rightarrow \quad G'([A]) \leq G'([B]) \tag{10}$$

holds. We also say that G is more informative than G' if, in addition to (10), there exist A, B such that $G'([A]) \leq G'([B])$ but $G([A]) \not\leq G([B])$.

Our earlier example shows that the two quadratic maps given in (7) and (9) are not related in the sense of Def. 4.7, but that they are both more informative than the trivial map related to $\Phi(A) = \mathrm{tr}\,[A]$. The content of the next result is that the Fisher information map is distinguished among quadratic order morphisms with $\tau(E) = \mathrm{tr}\,[E]$, in the sense that it is the most informative.

Proposition 4.8. *Among quadratic order preserving maps with* $\tau(E) = \mathrm{tr}\,[E]$, *the Fisher information map* F *is the most informative; it is at least as informative than any other order preserving quadratic map.*

Proof. A quadratic order morphism g corresponds to a linear map Φ of the form

$$\Phi(E) = \sum_{i=1}^{k} \langle \alpha_i, E \rangle \beta_i,$$

where α_i is a linear operator on $\mathbb{C}^d$, $\langle \cdot, \cdot \rangle$ is the Hilbert-Schmidt inner product ($\langle A, B \rangle = \langle A | B \rangle$) and $\beta_i \in M_{D \times r}$. We thus have

$$\Phi(E)\Phi(E)^* = J|E\rangle\langle E|J^* \tag{11}$$

for $J = \sum_{i=1}^{k} \beta_i \otimes \langle \alpha_i |$. This shows that any such map is a post-processing (by conjugating with the J matrix) of the Fisher information map. Denoting by F the map that corresponds $\Phi(E) = |E\rangle$, we have from (11), for any measurement A,

$$G(A) = JF(A)J^*. \tag{12}$$

Hence, $F(A) \leq F(B)$ implies $G(A) \leq G(B)$. $\qquad\square$

We note that the proof of Prop. 4.8 shows actually even more than is stated. Namely, by (12) we can get any other quadratic order preserving maps from the Fisher information map, assuming that we use the same map τ in the denominator. For instance, the matrices J allowing to obtain the quadratic maps from Eqs. (7), (9) are given by:

$$J_{(7)} = \sum_{i=1}^{d} |i\rangle\langle iii|$$

$$J_{(9)} = \sum_{i,j=1}^{d} |i\rangle\langle ijj| = I_d \otimes \sum_{j=1}^{d} \langle jj|.$$
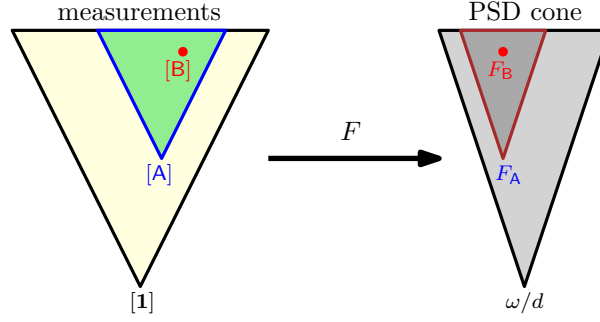
Figure 5: The Fisher information map F is an order morphism from the set of quantum measurements endowed with the post-processing order to the set of positive semidefinite matrices. We denote $F_X := F([X])$ for $X = \mathsf{A}, \mathsf{B}$.

In the formulas above, the product $J|E\rangle$ is understood with the ket $|E\rangle$ acting on the first two Hilbert space factors. Moreover note that the map from Eq. (9) is the partial trace of the Fisher information map with respect to the second subsystem. This conclusion naturally motivates to concentrate on the Fisher information map and that is what we do in the following sections.

5 The generalized Fisher information map

In [Zhu15, ZHC16] Huangjun Zhu and his collaborators considered the (classical) Fisher information matrix associated to a measurement A and to a statistical model centered at a given density matrix ρ. While in [Zhu15] the focus was on the maximally mixed state $\rho = \mathbb{1}/d$, the case of a general quantum state ρ was discussed at length in [ZHC16]. In this section, we view these constructions as order preserving maps from the set of measurements endowed with the post-processing partial order to the set of positive semidefinite matrices. We restate some of Zhu's results in the framework of Section 4, and we further develop the theory, adding new results and generalizations to this framework.

5.1 Properties of the Fisher information map

Given ρ a (fixed) faithful quantum state, that is $\rho > 0$, one can consider a generalized version of the Fisher information map induced by

$$f_\rho(E) = \frac{|\rho^{1/2}E\rangle\langle E\rho^{1/2}|}{\text{tr}\,[\rho E]}. \tag{13}$$

In the case of the (un-normalized) maximally mixed state, we simply write (see Figure 5 for a graphical representation):

$$f(E) = \frac{|E\rangle\langle E|}{\text{tr}\,[E]}. \tag{14}$$

We denote by F_ρ and F the maps that correspond to f_ρ and f, respectively, via (4).

Most of the following important results are contained in [Zhu15, ZHC16]. The new result here is the upper bound $\text{tr}\,[F([\mathsf{A}])] \leq \ell([\mathsf{A}])$, where $\ell([\mathsf{A}])$ is the number of elements in the simple representative of the equivalence class; see Sec. 3. For the sake of completeness, we prove all inequalities and also characterize the cases where the equality holds.

Proposition 5.1. *For any equivalence class* $[\mathsf{A}] \in \mathcal{O}_d^{\sim}$ *and for any given faithful state* $\rho > 0$*, we have*

$$1 \le \operatorname{tr}[F_\rho([\mathsf{A}])] \le \min(d, \ell([\mathsf{A}])).$$

The equality cases are as follows:

- $\operatorname{tr}[F_\rho([\mathsf{A}])] = 1$ *iff* $[\mathsf{A}]$ *is the least element in* $\mathcal{O}_d^{\sim}$*, i.e., the class of trivial measurement.*

- $\operatorname{tr}[F_\rho([\mathsf{A}])] = d$ *iff* $[\mathsf{A}]$ *is a maximal element in* $\mathcal{O}_d^{\sim}$*, i.e., a class consisting of rank-1 measurements*

- $\operatorname{tr}[F_\rho([\mathsf{A}])] = \ell([\mathsf{A}])$ *iff the simple representative of* $[\mathsf{A}]$ *is a sharp measurement, i.e., the operators in the range are projections.*

Proof. For the first upper bound, consider a representative $\mathsf{A} \in [\mathsf{A}]$, with $\mathsf{A} = (\mathsf{A}_x)_{x=1}^n$. The effects A_x are positive semidefinite as well as the faithful state $\rho > 0$, hence, using the sub-multiplicativity of the trace, it holds that $\operatorname{tr}[\rho \mathsf{A}_x^2] \le \operatorname{tr}[\rho \mathsf{A}_x] \operatorname{tr}[\mathsf{A}_x]$. This implies

$$\operatorname{tr}[F_\rho(\mathsf{A})] = \sum_{x=1}^n \frac{\operatorname{tr}[\rho \mathsf{A}_x^2]}{\operatorname{tr}[\rho \mathsf{A}_x]} \le \sum_{x=1}^n \frac{\operatorname{tr}[\rho \mathsf{A}_x] \operatorname{tr}[\mathsf{A}_x]}{\operatorname{tr}[\rho \mathsf{A}_x]} = \sum_{x=1}^n \operatorname{tr}[\mathsf{A}_x] = \operatorname{tr}[I_{\mathcal{H}}] = d \,.$$

The equality $\operatorname{tr}[\rho \mathsf{A}_x^2] = \operatorname{tr}[\rho \mathsf{A}_x] \operatorname{tr}[\mathsf{A}_x]$ holds if and only if A_x has rank one.

For the second upper bound, consider a simple representative $\mathsf{A} \in [\mathsf{A}]$ (see Subsec. 3.3). Since A_x are effects, we have $\rho^{1/2} \mathsf{A}_x^2 \rho^{1/2} \le \rho^{1/2} \mathsf{A}_x \rho^{1/2}$. Therefore,

$$\operatorname{tr}[F_\rho(\mathsf{A})] = \sum_{x=1}^{\ell(\mathcal{A})} \frac{\operatorname{tr}[\rho \mathsf{A}_x^2]}{\operatorname{tr}[\rho \mathsf{A}_x]} \le \sum_{x=1}^{\ell(\mathcal{A})} \frac{\operatorname{tr}[\rho \mathsf{A}_x]}{\operatorname{tr}[\rho \mathsf{A}_x]} = \ell(\mathcal{A}) \,.$$

The equality takes place if and only if for all x and given $\rho > 0$, it holds $\operatorname{tr}[\rho \mathsf{A}_x^2] = \operatorname{tr}[\rho \mathsf{A}_x]$ which is equivalent to the A_x being orthogonal projections.

Let us now turn to the lower bound. For a representative $\mathsf{A} \in [\mathsf{A}]$ and given $\rho > 0$, by the Cauchy-Schwarz inequality, we have, for all x,

$$(\operatorname{tr}[\rho \mathsf{A}_x])^2 = \left(\operatorname{tr}\left[\rho^{1/2} (\rho^{1/2} \mathsf{A}_x) \right] \right)^2 \le \operatorname{tr}[\rho] \cdot \operatorname{tr}\left[(\rho^{1/2} \mathsf{A}_x)^* \rho^{1/2} \mathsf{A}_x \right] = \operatorname{tr}\left[\rho \mathsf{A}_x^2 \right] \,.$$

This implies

$$\operatorname{tr}[F_\rho(\mathsf{A})] = \sum_{x=1}^n \frac{\operatorname{tr}[\rho \mathsf{A}_x^2]}{\operatorname{tr}[\rho \mathsf{A}_x]} \ge \sum_{x=1}^n \frac{(\operatorname{tr}[\rho \mathsf{A}_x])^2}{\operatorname{tr}[\rho \mathsf{A}_x]} = \sum_{x=1}^n \operatorname{tr}[\rho \mathsf{A}_x] = 1 \,.$$

Equality in the Cauchy-Schwarz inequality above holds iff each effect is a multiple of the identity, i.e. the measurement A is trivial. $\square$

For $\rho = \mathbb{1}/d$, the inequality $\operatorname{tr}[F(\mathsf{A})] \le d$ is the well-known Gill-Massar inequality [GM00] for the d-dimensional Fisher information matrix. This was the original approach of [Zhu15]. Let us also note that, in formula (13), one can have zero denominators if $\operatorname{tr}[\rho \mathsf{A}_x] = 0$. This can only occur if the positive semidefinite operators ρ and A_x have orthogonal supports, in which case the numerator is also null. Hence, one can extend the framework by allowing non-invertible matrices ρ simply by omitting terms with null denominators from the sum. For the sake of simplicity, we shall only consider invertible matrices ρ in the current paper. We also note that using rank-one operator ρ is useless for

incompatibility detection. Namely, let us now consider a rank-one operator $\rho = |\psi\rangle\langle\psi|$, where $\psi \in \mathcal{H}$ is a unit vector. The generalized Fisher information map reads in this case

$$F_{|\psi\rangle\langle\psi|}(\mathsf{A}) = \sum_x \frac{\left||\psi\rangle\langle\psi|\mathsf{A}_x\right\rangle\left\langle|\psi\rangle\langle\psi|\mathsf{A}_x\right|}{\langle\psi|\mathsf{A}_x|\psi\rangle} = |\psi\rangle\langle\psi| \otimes \sum_x \frac{\mathsf{A}_x^\top|\bar\psi\rangle\langle\bar\psi|\overline{\mathsf{A}_x}}{\langle\psi|\mathsf{A}_x|\psi\rangle} = |\psi\rangle\langle\psi| \otimes G_{\bar\psi}(\mathsf{A}^\top),$$

where $G_{\bar\psi}$ is the map from Eq. (6) and the complex conjugation is taken in the canonical basis. In particular, such a map cannot detect incompatibility, so the generalized Fisher information map with pure ρ is useless as an incompatibility criterion.

We note that one can consider the "truncated" map

$$\bar{F}_\rho(\mathsf{A}) = \sum_x \frac{1}{\operatorname{tr}[\rho\mathsf{A}_x]} |\rho^{1/2}\bar{\mathsf{A}}_x\rangle\langle\bar{\mathsf{A}}_x\rho^{1/2}|, \tag{15}$$

where $\bar{\mathsf{A}}_x := \mathsf{A}_x - \operatorname{tr}[\rho\mathsf{A}_x]\mathbb{1}$ is traceless with respect to ρ, i.e. $\operatorname{tr}\left[\rho\bar{\mathsf{A}}_x\right] = 0$. It holds that

$$F_\rho(\mathsf{A}) - \bar{F}_\rho(\mathsf{A}) = (\rho^{1/2}\otimes\mathbb{1})\omega(\rho^{1/2}\otimes\mathbb{1}), \tag{16}$$

where ω is the un-normalized maximally entangled density matrix $\omega = |\Omega\rangle\langle\Omega|$ (or the identity supe-operator). Working with this version of the map can have the benefit that trivial measurements are mapped to the zero matrix. However, for the sake of simplicity, we shall focus in this work on the "full" version of Fisher information map from (13) which is such that, for any trivial measurement T, we have

$$F_\rho(\mathsf{T}) = (\rho^{1/2}\otimes\mathbb{1})\omega(\rho^{1/2}\otimes\mathbb{1}).$$

In the following we discuss two important properties of the Fisher information map; the first is related to combining of two measurements on separate systems while the second one is related to adding noise to a measurement.

Proposition 5.2. *The map F_ρ respects tensor products: for any positive semidefinite matrices $\rho \in \mathsf{M}_d(\mathsf{C})$, $\sigma \in \mathsf{M}_{d'}(\mathsf{C})$, and measurements A, B of appropriate dimensions, we have*

$$F_{\rho\otimes\sigma}(\mathsf{A}\otimes\mathsf{B}) = F_\rho(\mathsf{A}) \otimes F_\sigma(\mathsf{B}),$$

where $F_{\rho\otimes\sigma}(\mathsf{A}\otimes\mathsf{B}) := \sum_{x,y} \frac{|\sqrt{\rho\otimes\sigma}\mathsf{A}_x\otimes\mathsf{B}_y\rangle\langle\mathsf{A}_x\otimes\mathsf{B}_y\sqrt{\rho\otimes\sigma}|}{\operatorname{tr}[(\rho\otimes\sigma)(\mathsf{A}_x\otimes\mathsf{B}_y)]}.$

Proof. The result follows from a simple computation, using the factorization property of the vectorization operation:

$$|X \otimes Y\rangle = |X\rangle \otimes |Y\rangle.$$

$\square$

A noisy measurement is typically described as a mixture of the original measurement with some trivial measurement [DFK19]. We remark that the mixture defined in the next statement is different than the one used in Prop. 4.5, where the measurements are being concatenated.

Proposition 5.3. *For any measurement A and every $\lambda \in [0,1]$, we denote*

$$\mathsf{A}_i^{(\lambda)} := \lambda\mathsf{A}_i + (1-\lambda)\operatorname{tr}[\rho\mathsf{A}_i]\mathbb{1}.$$

Then, for any given faithful state $\rho > 0$, we have

$$F_\rho(\mathsf{A}^{(\lambda)}) = \lambda^2 F_\rho(\mathsf{A}) + (1-\lambda^2)\rho^{1/2}\omega\rho^{1/2}, \tag{17}$$

and

$$\overline{F_\rho}(\mathsf{A}^{(\lambda)}) = \lambda^2\overline{F_\rho}(\mathsf{A}). \tag{18}$$

Proof. These are direct calculations. □

Especially (18) makes it effortless to study the effect of this kind of noise to measurements and their incompatibility.

5.2 Fisher information map on specific classes of measurements

The Fisher information map has a particular form for sharp measurements, making it suitable for analysis.

Proposition 5.4. *If* A *is a sharp measurement, then both* $F(\mathsf{A})$ *and* $\bar{F}(\mathsf{A})$ *are orthogonal projections of respective ranks* $\ell(\mathsf{A})$ *and* $\ell(\mathsf{A}) - 1$.

Proof. The equality $F(\mathsf{A})^2 = F(\mathsf{A})$ follows by a direct computation. The maps $F(\mathsf{A})$ and $\bar{F}(\mathsf{A})$ are connected by (16). Therefore, it is straightforward to see that $\bar{F}(\mathsf{A})$ is also a projection and, by (5.1), has the rank one less than that of $F(\mathsf{A})$. □

The Fisher information map is not injective, hence it does not preserve the full order structure of the triangle. However, in the following we show that it is one-to-one with respect to the subset of sharp measurements (i.e. those corresponding to projective POVMs) and the subset of dichotomic measurements.

Proposition 5.5. *Let* A *and* B *be two* sharp *measurements such that* $F(\mathsf{A}) = F(\mathsf{B})$. *Then* A *and* B *are post-processing equivalent:* $[\mathsf{A}] = [\mathsf{B}]$.

Proof. First, using the trace of the operator $F(\mathsf{A}) = F(\mathsf{B})$ and Prop. 5.1, we conclude that the measurements A, B have the same number of outcomes. Let this be k. Let $\{a_1, \ldots, a_d\}$ be an orthonormal basis of $\mathbb{C}^d$ in which the projections A_x are diagonal. By a direct computation, we have

$$\sum_{i=1}^{d} \langle a_i \otimes \bar{a}_i | F(\mathsf{A}) | a_i \otimes \bar{a}_i \rangle = k \, .$$

We then have

$$k = \sum_{i=1}^{d} \langle a_i \otimes \bar{a}_i | F(\mathsf{B}) | a_i \otimes \bar{a}_i \rangle = \sum_{i=1}^{d}\sum_{j=1}^{k} \frac{\langle a_i|\mathsf{B}_j|a_i\rangle^2}{\operatorname{tr}[\mathsf{B}_j]} \leq \sum_{i=1}^{d}\sum_{j=1}^{k} \frac{\langle a_i|\mathsf{B}_j|a_i\rangle}{\operatorname{tr}[\mathsf{B}_j]} = \sum_{j=1}^{k} \frac{\operatorname{tr}[\mathsf{B}_j]}{\operatorname{tr}[\mathsf{B}_j]} = k.$$

Since the equality holds above, we have that, for all i, j, $\langle a_i|\mathsf{B}_j|a_i\rangle \in \{0, 1\}$, showing that the projections B_j are also diagonal in the basis $\{a_1, \ldots, a_d\}$. In turn, this shows that the projections of A form a sub-partition of the projections of B. We can then have the same argument by replacing the roles of A and B, hence the projections of B form a sub-partition of the projections of A. Therefore, A and B are post-processing equivalent. □

Example 5.6. For a von Neumann measurement A, defined as $\mathsf{A}_i = |e_i\rangle\langle e_i|$ for a basis $\{e_i\}$, we have

$$F_\rho(\mathsf{A}) = \sum_{i=1}^{d} \frac{[(\rho^{1/2}e_i) \otimes \bar{e}_i][e_i \otimes \overline{(e_i\rho^{1/2})}]^*}{\operatorname{tr}[\rho e_i e_i^*]} = \sum_{i=1}^{d} \frac{\rho^{1/2}|e_i\rangle\langle e_i| \otimes \overline{|e_i\rangle\langle e_i|}\rho^{1/2}}{\operatorname{tr}[\rho|e_i\rangle\langle e_i|]} \, .$$

We can also show that the Fisher information map is injective on dichotomic measurements. This is a simple consequence of the following expression: for a dichotomic measurement A consisting of effects E and $\mathbb{1} - E$, we have

$$F_\rho(\mathsf{A}) = \rho^{1/2}\left[|\alpha\mathbb{1} + \beta E\rangle\langle\alpha\mathbb{1} + \beta E| + \omega/\mathrm{tr}\,[\rho]\right]\rho^{1/2},$$

where

$$\alpha = \sqrt{\frac{\mathrm{tr}\,[\rho E]}{\mathrm{tr}\,[\rho]\,\mathrm{tr}\,[\rho(\mathbb{1} - E)]}} \quad\text{and}\quad \beta = -\sqrt{\frac{\mathrm{tr}\,[\rho]}{\mathrm{tr}\,[\rho E]\,\mathrm{tr}\,[\rho(\mathbb{1} - E)]}}.$$

In the case where $\rho = \mathbb{1}$, we get

$$\alpha = \sqrt{\frac{\mathrm{tr}\,[E]}{d(d - \mathrm{tr}\,[E])}} \quad\text{and}\quad \beta = -\sqrt{\frac{d}{\mathrm{tr}\,[E]\,(d - \mathrm{tr}\,[E])}}.$$

Hence, one can recover the operator E from the knowledge of $F_\rho(\mathsf{A})$: from the knowledge of ρ, $F_\rho(\mathsf{A})$ and the fact that $\mathbb{1}$ and E are independent (otherwise E is a trivial effect), one can recover $|E\rangle\langle E|$ and then, up to a phase, E.

We will then discuss an important class of measurements for which the Fisher information map is not injective. We start by recalling that, as was pointed out in [ZE11], the non-singularity of the image of the map F is related to the measurement being informationally complete (i.e. the linear span of $\{\mathsf{A}_i\}_i$ is $\mathcal{L}(\mathcal{H})$).

Proposition 5.7. *Let $\rho > 0$. A measurement A is informationally complete iff the matrix $F_\rho(\mathsf{A})$ is invertible.*

Proof. It is a well-known fact (see e.g. [Sco06]) that A is informationally complete iff the frame operator

$$\mathcal{F}_\mathsf{A} = \sum_x |\mathsf{A}_x\rangle\langle\mathsf{A}_x|$$

is invertible. This happens iff the $d^2 \times d^2$ matrix $F_\rho(\mathsf{A})$ is invertible, since all the denominators $\mathrm{tr}\,[\rho\mathsf{A}_x]$ are positive. $\qquad\square$

For an arbitrary (i.e. possibly continuous) measurement C, we define the Fisher information map as follows:

$$F(\mathsf{C}) = \int_\Omega \frac{|\mathsf{C}(\omega)\rangle\langle\mathsf{C}(\omega)|}{\mathrm{tr}\,[\mathsf{C}(\omega)]}\,\mathrm{d}\omega.$$

Example 5.8. For the continuous POVM $\mathsf{C} = d\{|\psi\rangle\langle\psi|\}$, where ψ is a uniformly distributed point on the unit sphere of $\mathbb{C}^d$, we have

$$F(\mathsf{C}) = \frac{2}{d+1}P_{sym}^\Gamma,$$

where P_{sym} is the orthogonal projection on the symmetric subspace of $\mathbb{C}^d \otimes \mathbb{C}^d$, and the Γ symbol denotes the partial transposition. Indeed, using $||\psi\rangle\langle\psi|\rangle = |\psi\rangle \otimes \overline{|\psi\rangle}$, we have

$$F(\mathsf{C}) = d\int_{\|\psi\|=1} |\psi\rangle\langle\psi| \otimes \overline{|\psi\rangle\langle\psi|}\,\mathrm{d}\psi = d\left[\int_{\|\psi\|=1} |\psi\rangle\langle\psi| \otimes |\psi\rangle\langle\psi|\,\mathrm{d}\psi\right]^\Gamma \tag{19}$$

$$= d\frac{P_{sym}^\Gamma}{d(d+1)/2} = \frac{2}{d+1}P_{sym}^\Gamma. \tag{20}$$

The third equality above follows from [Har13, Proposition 2]. Note that above, $\mathrm{d}\psi$ denotes the *normalized* Lebesgue measure on the unit sphere of $\mathbb{C}^d$, in such a way that

$$\int |\psi\rangle\langle\psi|\,\mathrm{d}\psi = \frac{\mathbb{1}}{d}.$$

Proposition 5.9. *Let μ be a complex spherical 2-design, that is μ is a probability measure on the unit sphere of $\mathbb{C}^d$ such that*

$$\binom{n+d-1}{n} \int |\psi\rangle\langle\psi|^{\otimes n} \mathrm{d}\mu(\psi) = P_{sym}^{(n)} \qquad \text{for } n = 1, 2, \tag{21}$$

where $P_{sym}^{(n)}$ is the orthogonal projection on the symmetric subspace of $(\mathbb{C}^d)^{\otimes n}$. Then, the POVM $\mathsf{A} = d\{\mu(\psi)|\psi\rangle\langle\psi|\}$ is such that

$$F(\mathsf{A}) = \frac{2}{d+1}[P_{sym}^{(2)}]^{\Gamma}.$$

Proof. First, note that (21) for $n = 1$ is precisely the condition that A is a POVM. For $n = 2$, equation (19) still holds when replacing the Lebesgue measure $\mathrm{d}\psi$ by the 2-design $\mathrm{d}\mu(\psi)$. $\qquad\square$

Corollary 5.10. *The Fisher information map map F is not injective, since any 2-design is mapped to $2P_{sym}/(d+1)$.*

Remark 5.11. Let A be a rank-1 measurement, i.e., a maximal element $\mathcal{O}_{\tilde{d}}$. Then

$$F(\mathsf{A})^{\Gamma} \in \mathrm{conv}\{|a\rangle\langle a|^{\otimes 2} \,:\, a \in \mathbb{C}^d\}.$$

In particular, $F(\mathsf{A})^{\Gamma}$ is a separable, symmetric, positive semidefinite matrix.

5.3 Measurement incompatibility criterion

Up to this point, we have focused on the properties of the Fisher information maps relative to the partial order of post-processing. The paramount application of this map is to detect incompatibility of measurements and this is the topic of the current section. We gather here the main theoretical results and we present applications in the next section.

The incompatibility criterion for quantum measurements that we obtain can be expressed with the help of the following height function for matrices.

Definition 5.12. For a finite set of $d \times d$ self-adjoint matrices $X = \{X_i\}_{i \in [n]}$, define the *height* of X as

$$\mathfrak{h}(X) := \min\{\mathrm{tr}\,[H] \,:\, H \geq X_i \quad \forall i \in [n]\}. \tag{22}$$

In Figure 6 we provide an interpretation of the height function: it is the minimum trace of an operator H which dominates, in the PSD order, the set $\{X_1, \dots, X_n\}$. Since the positive semidefinite (or Löwner) cone is not a lattice, the set of matrices H above has a complicated structure (e.g. it does not admit a least element in general), and thus the minimization problem is not trivial. In some cases height function is easy to calculate. For instance, if the matrices in X are commuting, they can be diagonalized in some common basis, and one can easily find the smallest (entry-wise) diagonal matrix H dominating all the X_i.

A crucial fact about the height function is that it can be computed with the help of a semidefinite program (SDP) [BV04]. In the case of two matrices, one can get an explicit formula for $\mathfrak{h}$, see [HJN20, Eqs. (13)-(15)]:

$$\mathfrak{h}(X_1, X_2) = \frac{1}{2}\left(\mathrm{tr}\,[X_1] + \mathrm{tr}\,[X_2] + \|X_1 - X_2\|_1\right), \tag{23}$$

where $\|\cdot\|_1$ denotes the Schatten 1-norm (or the nuclear norm).

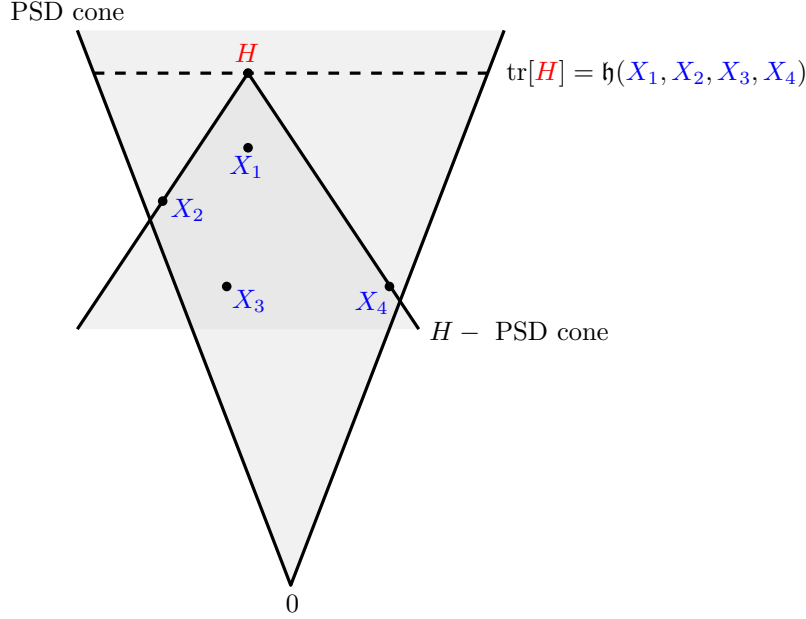In the general setting, the height function satisfies the following set of properties.

Figure 6: The height function of a set of four matrices $X_{1,2,3,4}$. Note that in general, the set $\{X_i\}$ does not have a smallest upper bound in the PSD order.

Proposition 5.13. *Given positive semidefinite matrices $X_1, \ldots, X_n$, the following upper bound holds:*

$$\mathfrak{h}(X) \leq \sum_{i=1}^{n} \operatorname{tr}\left[X_i\right].$$

Moreover, if the matrices X_i have orthogonal supports, the inequality above is saturated.

Proof. The first claim follows from the fact that $H = \sum_{i=1}^{n} X_i$ is feasible point of the semidefinite program (22) and hence, $\operatorname{tr}\left[H\right] \geq \operatorname{tr}\left[\sum_{i=1}^{n} X_i\right]$. The second claim follows from the following implication, which is true if the matrices X_i have orthogonal supports:

$$\forall i \in [n], \quad H \geq X_i \implies H \geq \sum_{i=1}^{n} X_i.$$

$\square$

One can compute the dual of the SDP in Eq. (22) [BV04, Chapter 5] and obtain:

$$\mathfrak{h}(X) = \max\left\{\sum_{i=1}^{n} \operatorname{tr}\left[X_i Y_i\right] \,:\, Y = (Y_1, \ldots, Y_n) \text{ is a } n\text{-outcome measurement}\right\}. \quad (24)$$

In particular, any choice of a measurement Y gives a *lower bound* on the quantity $\mathfrak{h}(X)$. Moreover, in the case when the matrices X_i are positive semidefinite, an interesting choice is the *pretty-good-measurement* [HW94]:

$$Y_i = S^{-1/2} X_i S^{-1/2}, \qquad \text{with } S = \sum_{i=1}^{n} X_i,$$

where one uses the pseudo-inverse to compute the inverse matrix square roots.

The following result is the *incompatibility criterion* discovered by Zhu [Zhu15]. We slightly improve it below, by adding another constraint depending on the number of measurement outcomes.

Theorem 5.14. *Let $\mathcal{A} = \{A^{(1)}, \ldots, A^{(g)}\}$ be a set of g d-dimensional measurements. For any faithful quantum state ρ, we denote*

$$F_\rho(\mathcal{A}) := \left\{ F_\rho(A^{(i)}) \right\}_{i=1}^g.$$

If there exists $\rho > 0$ such that

$$\mathfrak{h}(F_\rho(\mathcal{A})) > \min\left(d, \prod_{i=1}^g \ell(A^{(i)})\right)$$

then $\mathcal{A}$ is incompatible.

Proof. Assuming that the given measurements are compatible, let B be their joint measurement. Since every $A^{(i)}$ is a post-processing of B, it follows from Prop. 4.3 that $F_\rho(A^{(i)}) \leq F_\rho(B)$ for all $i \in [g]$ and thus, using (22), we get $\mathfrak{h}(F_\rho(\mathcal{A})) \leq \operatorname{tr}[F_\rho(B)]$. From Prop. 5.1 we know that

$$\operatorname{tr}[F_\rho(B)] \leq \min(d, \ell(B))$$

and the conclusion follows from the fact that one can always choose a joint measurement having $\prod_{i=1}^g \ell(A^{(i)})$ outcomes. $\qquad\square$

Remark 5.15. If the measurements $A^{(1)}, \ldots, A^{(g)}$ are such that

$$\sum_{i=1}^g \operatorname{tr}\left[F_\rho(A^{(i)})\right] \leq \min\left(d, \prod_{i=1}^g \ell(A^{(i)})\right),$$

it follows from Proposition 5.13 that one cannot witness the incompatibility of these measurements using the criterion from Theorem 5.14.

Using the dual formulation of the SDP used for defining the quantity $\mathfrak{h}$ and the pretty-good-measurement of [HW94], we obtain the following simpler incompatibility criterion, which does not require computing the value of a semidefinite program.

Corollary 5.16. *Let $\mathcal{A} = \{A^{(1)}, \ldots, A^{(g)}\}$ be a set of g d-dimensional measurements. We define a measurement Y with g outcomes by*

$$Y_i = S^{-1/2} F_\rho(A^{(i)}) S^{-1/2}, \qquad \text{with } S = \sum_{i=1}^g F_\rho(A^{(i)}).$$

If

$$\sum_{i=1}^g \operatorname{tr}\left[F_\rho(A^{(i)}) Y_i\right] > \min\left(d, \prod_{i=1}^g \ell(A^{(i)})\right),$$

then $\mathcal{A}$ is incompatible.

Remark 5.17. In the dual optimization problem from eq. (24), one can recognize an expression similar to the one of the optimal zero-error, one-shot, distinguishability of n quantum states of dimension d^2. Indeed, the function $\mathfrak{h}$ is a measure of distinguishability of the matrices X_i. Hence, the more distinguishable the matrices X_i are, the more incompatible they are (in the sense of the criterion from Theorem 5.14).

One can use the newly introduced bound from Proposition 5.1 to give a lower bound on the number of outcomes of any joint measurement for a g-tuple of compatible measurements.

Proposition 5.18. *Let $\mathcal{A} = \{A^{(1)}, \dots, A^{(g)}\}$. For any faithful quantum state $\rho > 0$, the length of any joint measurement B (i.e. $B \succcurlyeq A^{(i)}$ for all i) satisfies*

$$\ell(B) \geq \mathfrak{h}(F_\rho(\mathcal{A})).$$

Proof. The proof is a simple consequence of Proposition 4.3 and Proposition 5.1 (or directly from Theorem 5.14). $\qquad\square$

Remark 5.19. We recall that it has been shown in [SHSL20] that, given a g-tuple of compatible measurements, one can always find a joint measurement with a number of outcomes which is *linear* in the number g of measurements. This is an interesting result, but it does not yield a better incompatibility criterion in Theorem 5.14 since the value is always larger than d.

6 Examples

In this section we investigate the efficiency of our approach in some concrete examples, for which we can find the incompatibility conditions for the given measurements and compare them to known necessary and sufficient conditions. As we will demonstrate, the Fisher information incompatibility condition is exact in some cases, good in some cases, while relatively poor in some cases.

6.1 Fourier-conjugate bases

The compatibility problem for noisy versions of Fourier-conjugate bases has been studied previously in [CHT12, CHT19]. We recall that two d-outcome measurements A and B, are called the noisy version of Fourier-conjugate measurements if

$$A_i = s|e_i\rangle\langle e_i| + (1 - s)\frac{\mathbb{1}}{d}, \qquad s \in [0, 1]$$

$$B_j = t\mathcal{F}|e_j\rangle\langle e_j|\mathcal{F}^* + (1 - t)\frac{\mathbb{1}}{d}, \qquad t \in [0, 1]$$

where $\mathcal{F}$ is the Fourier matrix $\mathcal{F}_{ab} = \zeta^{ab}/\sqrt{d}$ with $\zeta = \exp(2\pi i/d)$ a primitive root of unity.

It has been shown in [CHT12, Prop. 6] that A and B are compatible iff

$$s + t \leq 1 \quad \text{or} \quad s^2 + t^2 + \frac{2(d-2)}{d}(1-s)(1-t) \leq 1. \tag{25}$$

Let us apply Fisher information incompatibility criterion from Section 5.3 in this situation. As it has been shown in [Zhu15, Equation (23)], a necessary condition for compatibility is $s^2 + t^2 \leq 1$. The application of the criterion is straightforward, since for mutually unbiased bases (MUBs) A and B, the matrices $\bar{F}(A)$ and $\bar{F}(B)$ are orthogonal, and Prop. 5.13 applies. We notice that for $d = 2$ the two conditions of incompatibility, that are the conditions given by Theorem 5.14 and equation (25), are the same. It is of interest to compare the corresponding incompatibility regions of the parameters (s, t) for different dimension values and this is presented in Figure 7. As we can see, in general, the shaded region that corresponds to pairs (s, t) for which incompatibility is guaranteed by Zhu's Fisher information criterion is much smaller respect to the region given by (25).
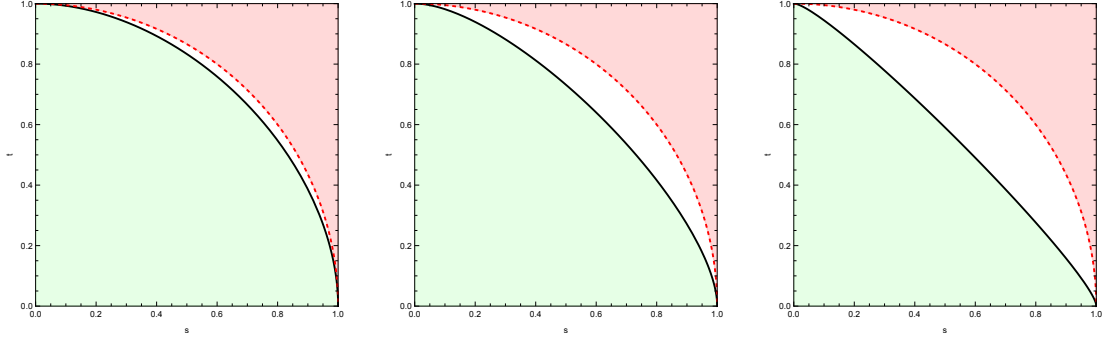
Figure 7: The region of parameters $(s, t) \in [0, 1]^2$ for which noisy versions of Fourier-conjugated bases are compatible: left $d = 3$, center $d = 10$, right $d = 100$. The compatibility region computed in [CHT12] is delimited by the thick black curve, with the green shaded region corresponding to pairs (s, t) of compatible measurements. The Fisher information incompatibility criterion corresponds to the thick dotted red curve. The red shaded region corresponds to pairs (s, t) for which incompatibility is guaranteed by Zhu's Fisher information criterion. For $d = 2$ the two curves overlap.

6.2 Two unbiased dichotomic qubit measurements

We consider now the important examples of unbiased dichotomic qubit measurements given by $\mathsf{A}^{(k)}(\pm 1) = \frac{1}{2}(\mathbb{1} \pm \eta_k \vec{n}_k \cdot \sigma)$, $k = 1, 2, 3$, where $0 \le \eta_k \le 1$ is the sharpness, $\vec{n}_k = \vec{e_k}$, $k = 1, 2, 3$ are the vectors of canonical basis in $\mathbb{R}^3$ and $\sigma = (\sigma_x, \sigma_y, \sigma_z)$ is the vector of the Paulli operators. It has been shown in [Bus86] that any two among these three measurements are compatible iff $\eta_k^2 + \eta_\ell^2 \le 1$.

We aim to compute in this case the Fisher information matrix

$$F_\rho(\mathsf{A}^{(k)}) := \sum_i \frac{|\rho^{1/2} A_i^{(k)}\rangle \langle \rho^{1/2} A_i^{(k)}|}{\operatorname{tr}\left[\rho A_i^{(k)}\right]},$$

for $\rho = \frac{1}{2} \begin{pmatrix} 1 + v_3 & v_1 - iv_2 \\ v_1 + iv_2 & 1 - v_3 \end{pmatrix}$ and $v = (v_1, v_2, v_3)^T, ||v|| \le 1$ is the Bloch vector of the state ρ.

To exemplify our theory we present the problem in the setting $v_1 = v_2 = 0$. In this case, the Fisher information matrices of the measurements $\mathsf{A}^{(k)}, k = 1, 2, 3$ can be written

explicitly as

$$F_\rho(\mathsf{A}^{(1)}) = \frac{1}{2}\begin{pmatrix} (1+v_3) & 0 & 0 & \sqrt{1-v_3^2} \\ 0 & \eta_1^2(1-v_3) & \eta_1^2\sqrt{1-v_3^2} & 0 \\ 0 & \eta_1^2\sqrt{1-v_3^2} & \eta_1^2(1+v_3) & 0 \\ \sqrt{1-v_3^2} & 0 & 0 & (1-v_3) \end{pmatrix}$$

$$F_\rho(\mathsf{A}^{(2)}) = \frac{1}{2}\begin{pmatrix} (1+v_3) & 0 & 0 & \sqrt{1-v_3^2} \\ 0 & \eta_2^2(1-v_3) & -\eta_2^2\sqrt{1-v_3^2} & 0 \\ 0 & -\eta_2^2\sqrt{1-v_3^2} & \eta_2^2(1+v_3) & 0 \\ \sqrt{1-v_3^2} & 0 & 0 & (1-v_3) \end{pmatrix}$$

$$F_\rho(\mathsf{A}^{(3)}) = \frac{1}{2}\begin{pmatrix} (1+v_3)(a^2+1) & 0 & 0 & \sqrt{1-v_3^2}(1-ab) \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \sqrt{1-v_3^2}(1-ab) & 0 & 0 & (1-v_3)(b^2+1) \end{pmatrix} \tag{26}$$

where $a = \frac{\eta_3 - \eta_3 v_3}{\sqrt{1-\eta_3^2 v_3^2}}$, $b = \frac{\eta_3 + \eta_3 v_3}{\sqrt{1-\eta_3^2 v_3^2}}$.

It follows that $\text{tr}\left[F_\rho(\mathsf{A}^{(1)})\right] = \eta_1^2 + 1$, $\text{tr}\left[F_\rho(\mathsf{A}^{(2)})\right] = \eta_2^2 + 1 + v_3$ and $\text{tr}\left[F_\rho(\mathsf{A}^{(3)}))\right] = 1 + \frac{\eta_3^2 - \eta_3^2 v_3^2}{1-\eta_1^2 v_3^2}$. We aim to compute the height of Fisher information matrices $F_\rho(\mathsf{A}^{(k)})$, $k = 1, 2, 3$, for finding the incompatibility region given by the Theorem 5.14. In the case of two measurements, for example, $\mathsf{A}^{(1)}$ and $\mathsf{A}^{(3)}$ (corresponding to σ_x and σ_z) we find that the incompatibility condition given by the SDP (23) is

$$\eta_1^2 + \eta_3^2 - \eta_1^2 \eta_3^2 v_3^2 > 1, \tag{27}$$

where $(\eta_1, \eta_3) \in [0,1]^2$.

We notice that for general $v_3(|v_3| \leq 1)$ we get a region of compatibility which is larger than the quatercircle $\eta_1^2 + \eta_3^2 \leq 1$, corresponding to the case $v_3 = 0$ ($\rho = \frac{1}{2}$). For $v_3 = \pm 1$ (corresponding to ρ pure state) we get the full square, so the criterion is uninformative. For $(\eta_1, \eta_3) = (1, 0)$, we get that the compatibility holds for any $v_3, |v_3| \leq 1$.

In the case of the two sharp measurements corresponding to the observables σ_x and σ_y, for any v_3, we get the incompatibility condition $\eta_1^2 + \eta_2^2 > 1$. This result suggests that considering a state ρ given by a Bloch vector of the form $v = (0, 0, v_3)$ gives the same incompatibility condition, independent of v_3.

It is of interest to study the influence of a generic state ρ in determining the incompatibility region for two unbiased dichotomic qubit measurements, for example, the cases corresponding to σ_x, σ_z and σ_x, σ_y. We are going to study this by computing the height function $\mathfrak{h}\left(F_\rho(\mathcal{A})\right) \equiv h(v_1, v_2, v_3, \eta_1, \eta_2, \eta_3)$ which is now a function of the sharpness $\eta_i, i = 1, 2, 3$ and of the parameters of the Bloch vector $v_i, i = 1, 2, 3$. (8)

As we can see from Figures 8 (the case σ_x and σ_y) and 9 (the case σ_x and σ_z), the limiting condition for incompatibility $h(v_1, v_2, v_3, \eta, \eta, \eta) > 2$ is better achieved using the Bloch vector $v = (0, 0, 0)$, that corresponds to the maximally mixed state, comparing to another specific Bloch vector, for example $v = (1/2, 1/3, 1/3)$. We also notice that using unit Bloch vector does not provide information about the region of incompatibility. Our graphics confirms the results that two unbiased dichotomic qubit measurements of the same sharpness η are incompatible for $\eta > 1/\sqrt{2} \approx 0.7$, which actually coincide with the value of $h(v_1, v_2, v_3, \eta, \eta, \eta) > 2$, for which it becomes grater that 2.

Accepted in ⟨ ⟩uantum 2022-10-22, click title to verify. Published under CC-BY 4.0.
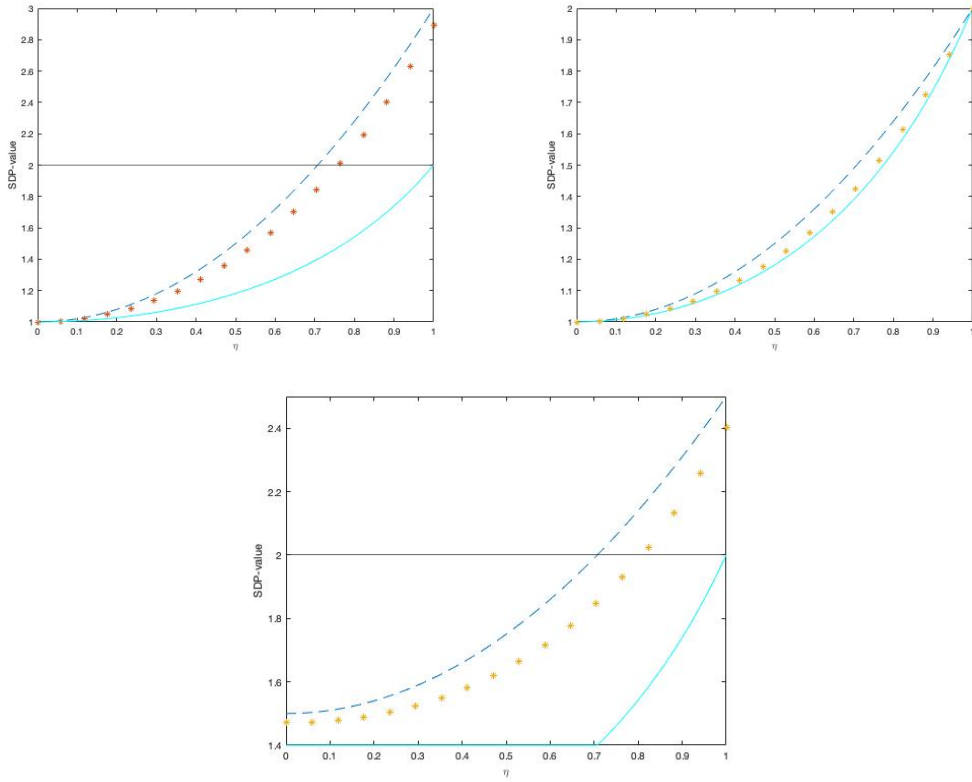
24

Figure 8: The plot of the height function $h(v_1, v_2, v_3, \eta, \eta, \eta)$ (first), $h(v_1, v_2, v_3, 0, \eta, 0)$ (second) and $h(v_1, v_2, v_3, 0, \eta, 1/\sqrt{2})$ (third) of the generalized Fisher information matrix corresponding to two unbiased dichotomic qubit measurements (σ_x, σ_y) for $\rho = \frac{1}{2}(\mathbb{1} + v \cdot \sigma)$; $v = (0,0,0)$ is dashed $(--)$ line, $v = (1/2, 1/3, 1/3, 0)$ is star $(*)$ line and $v = (1/\sqrt{3}, 1/\sqrt{3}, 1/\sqrt{3})$ for cyan line.
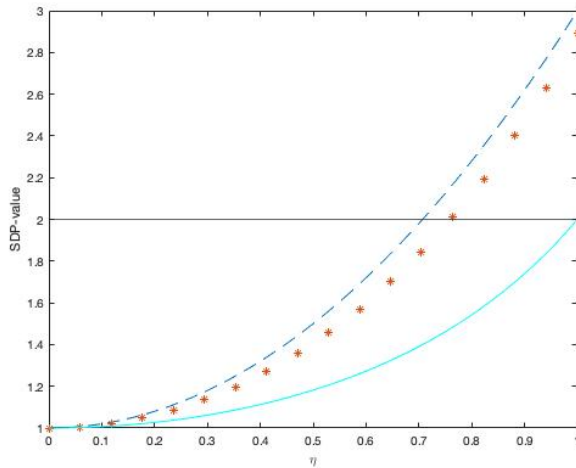
Figure 9: The plot of the height function $h(v_1, v_2, v_3, \eta, \eta, \eta)$ of the generalized Fisher information matrix corresponding to two unbiased dichotomic qubit measurements (σ_x, σ_z) for $\rho = \frac{1}{2}(\mathbb{1} + v \cdot \sigma)$; $v = (0,0,0)$ is dashed $(--)$ line, $v = (1/2, 1/3, 1/3)$ is star $(*)$ line and $v(1/\sqrt{3}, 1/\sqrt{3}, 1/\sqrt{3})$ for cyan line.

6.3 Unbiased quantum effects

Let $\mathsf{A}^{(i)}_\pm = ((\mathbb{1} \pm \mathsf{T}_i)/2)$ be a g-tuple of unbiased dichotomic measurements, for some traceless self-adjoint matrices $T_1, \ldots, T_g \in M_d(\mathbb{C})$. A simple computation shows that the truncated map $\bar{F}_i$ introduced in (15) can be written as $\bar{F}_i = |\mathsf{T}_i\rangle\langle\mathsf{T}_i|/d$. Assume now that all the T_i are unitary and anti-commuting; such a family exists, for example, if $d \geq 2^{g-1}$, see [New32] or [Hru16, Theorem 1]. It has been shown in [BN18, Theorems VII.7 and VIII.8] that such a g-tuple of quantum effects is maximally incompatible: the set of noise parameters $s \in [0,1]^g$ rendering the noisy measurements

$$\tilde{\mathsf{A}}^{(i)} = s_i \mathsf{A}^{(i)} + (1 - s_i)(\mathbb{1}/2, \mathbb{1}/2)$$

compatible is precisely the "quarter circle" $\{s \in [0,1]^g : \sum_{i=1}^g s_i^2 \leq 1\}$. The term *maximal incompatibility* refers to the fact that noise parameters inside the quarter circle generate compatible dichotomic measurements, for *any* starting measurements [BHSS13]. We analyze the next the power of Zhu's criterion in this setting. For unitary, self-adjoint, anti-commuting T_i's, we have

$$\mathrm{tr}\,[\mathsf{T}_i\mathsf{T}_j] = d\delta_{ij},$$

which implies that the matrices $\bar{F}_i$ are g orthonormal unit-rank projections. Zhu's criterion for the noisy versions reads

$$\sum_{i=1}^g s_i^2 \leq d - 1,$$

which is precisely the quarter circle condition in the case $d = 2$, and a much worse condition for larger d. For $d = 2$, this condition proves the optimality of $g = 2$ or $g = 3$ unbiased qubit effects; the matrices F are in this case the Pauli operators. The next cases, $g = 4, 5$, require $d = 4$ elements, the condition above still being non-trivial. However, for $g \geq 6$, we require $d \geq 8$, yielding the condition trivial, since

$$g \leq 2^{\lceil (g-1)/2 \rceil} - 1.$$

6.4 Planar qubit measurements

For the planar qubit effects defined in [ULMH16], after adding noise, the Zhu maps read

$$\bar{F}(a) = \lambda^2 |a\rangle\langle a|.$$

Requiring $H \geq \bar{F}(a)$ for $a = \cos(k\pi/M)e_1 + \sin(k\pi/M)e_2$ for $k = 0, 1, \ldots, M - 1$, the best choice is $H = \lambda^2 I_2$, which gives the inequality $\lambda \leq 1/\sqrt{2}$. From [ULMH16, AK20] we know that the optimal value is $\lambda \leq \frac{1}{M \sin(\pi/(2M))}$, so the Fisher information criterion is optimal iff $M = 2$.

6.5 Three unbiased dichotomic qubit measurements

We consider now the case of three unbiased dichotomic qubit measurements given by $\mathsf{A}_k(\pm 1) = \frac{1}{2}(\mathbb{1} \pm \eta_k \vec{n}_k \cdot \sigma)$, $k = 1, 2, 3$, where $\eta_k \leq 1$ is the sharpness. Necessary and sufficient condition for joint measurability of these measurements is presented in [YO13] in terms of Fermat-Torricelli (FT) point $\vec{v}_{FT}$ (the point in $\mathbb{R}^3$ that minimizes the sum of the distances from itself to a given set of points). It holds that the measurements A_k, $k = 1, 2, 3$, are triplewise jointly measurable if and only if

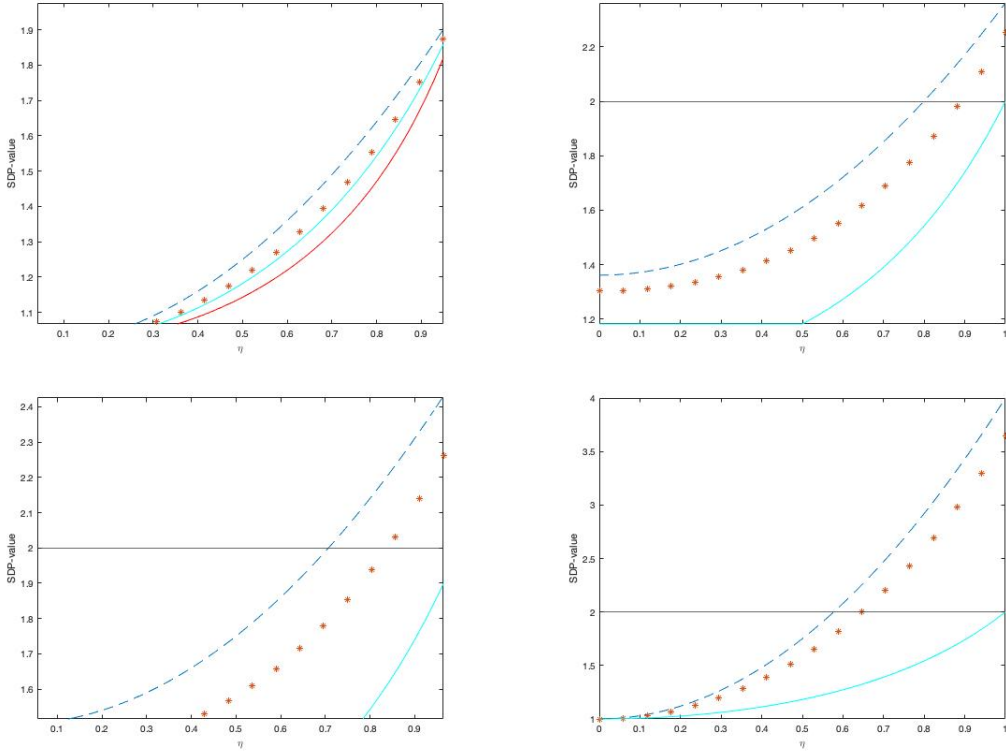$$\sum_{i=0}^3 \|\vec{v}_{FT} - \vec{v}_i\| \leq 4 \tag{28}$$

Figure 10: The plot of the height functions $h(v_1, v_2, v_3, \eta, 0, 0)$ (first), $h(v_1, v_2, v_3, \eta, 1/2, 1/3)$ (second) $h(v_1, v_2, v_3, \eta, 1/2, 1/2)$ (third) and $h(v_1, v_2, v_3, \eta, \eta, \eta)$ (fourth) of the generalized-Zhu matrix corresponding to three unbiased dichotomic qubit measurements (σ_x, σ_y, and σ_z) for $\rho = \frac{1}{2}(I + v \cdot \sigma)$; $v = (0,0,0)$ is dashed $(--)$ line, $v = (1/2, 1/3, 1/3, 0)$ is star $(*)$ line, $v = (1/\sqrt{2}, 0, 1/\sqrt{3})$ the read line and $v(1/\sqrt{3}, 1/\sqrt{3}, 1/\sqrt{3})$ for cyan line.
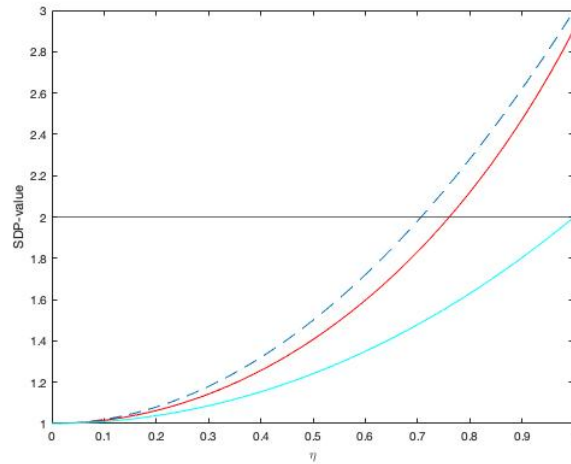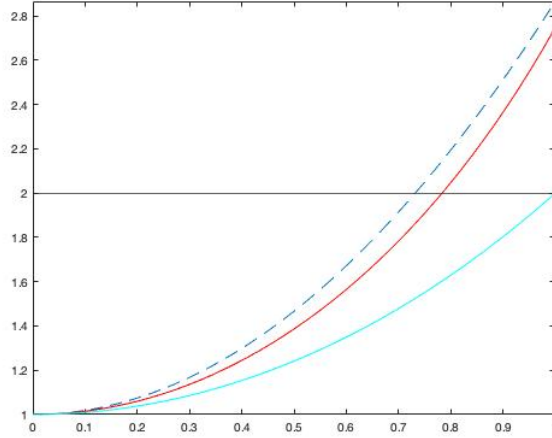

Figure 11: The plot of the height functions $h(v_1, v_2, v_3, \eta, \eta, \eta)$ of the generalized Fisher information map corresponding to three unbiased dichotomic qubit measurements corresponding to $\eta_1 = (0, 1, 0)$ and $\vec{n}_k = (\cos\frac{k-1}{3}\pi, \sin\frac{k-1}{3}\pi, 0), k = 2, 3$, for $\rho = \frac{1}{2}(\mathbb{1} + v \cdot \sigma)$; $v = (0,0,0)$ is dashed $(--)$ line, $v = (1/2, 1/3, 1/3)$ is red line

Figure 12: The plot of the height functions $h(v_1, v_2, v_3, \eta, \eta, \eta)$ of the generalized Fisher information map corresponding to three unbiased dichotomic qubit measurements corresponding to $\eta_1 = (0,0,1)$ and $\vec{n}_k = (\cos\frac{k-1}{3}\pi, \sin\frac{k-1}{3}\pi, 0), k = 2, 3$, for $\rho = \frac{1}{2}(I + v \cdot \sigma)$; $v = (0,0,0)$ is dashed ($--$) line, $v = (1/2, 1/3, 1/3)$ is red line.

where $\vec{v}_{FT}$ is FT point of the following four points in $\mathbb{R}^3$: $\vec{v}_0 = -\sum\limits_{k=1}^{3} \eta_k \vec{n}_k$ and $\vec{v}_k = -2\eta_k \vec{n}_k - \vec{v}_0, k = 1, 2, 3$. The FT point of four points does not have an analytical expression, except the case of coplanar vectors and the case when one of the vector is orthogonal to other two vectors.

In the following we make a comparison (see Table 1) between the necessary and sufficient compatibility condition given by the FT-point and Fisher information necessary incompatibility criterion, given by Theorem 5.14. We compare the compatibility region derived from the analytical condition of the FT point, with the one obtained by the Fisher information criterion (we use the Matlab `cvx` package for computing the height of Fisher information map, $\mathfrak{h}(F(\mathsf{A}_k))) \, k = 1, 2, 3$, introduced by eq. (22)). In the first line of the table (1) we present the case of trine spin measurements, that is the case of three measurements (of global sharpness η) in an equatorial plane of the Bloch ball equiangularly separated by an angle of $\pi/3$, that is the case $\vec{n}_k = (\cos\frac{k-1}{3}\pi, \sin\frac{k-1}{3}\pi, 0), k = 1, 2, 3$. The coplanar property of the vectors allows to derive explicitly the FT point and consequently, the necessary and sufficient condition for compatibility admits an analytical form [YO13, AK20]

$$\eta(\|\vec{n}_1 + \vec{n}_2\| + \|\vec{n}_1 - \vec{n}_3\| + \|\vec{n}_2 - \vec{n}_3\|) \leq 2 \tag{29}$$

We notice that Fisher information condition of incompatibility $\eta > \frac{1}{\sqrt{2}}$ is weaker by confront to the compatibility condition given by FT point, that is $\eta \leq 2/3$. This results confirms again the results presented at example (6.4), where we claim that Fisher information incompatibility condition is optimal for two planar orthogonal qubit measurements. Another interesting case is when one of the vectors, η_1, is orthogonal on the other two ($\vec{n}_1 \perp \vec{n}_2, \vec{n}_3$). More exactly, we set $\vec{n}_1 = (0,0,1), \vec{n}_k = (\cos\frac{k-1}{3}\pi, \sin\frac{k-1}{3}\pi, 0), k = 2, 3$. In this case, the FT point can be found explicitly and the corresponding compatibility condition (28) becomes:

$$\eta(\|\vec{n}_2 + \vec{n}_3\| + \|\vec{n}_2 - \vec{n}_3\|) \leq 2\sqrt{1 - \eta^2\|\vec{n}_1\|^2}, \tag{30}$$

and setting the compatibility region for $\eta \leq 0.59$. On the other hand Zhu incompatibility

condition requires $\eta > 0.72$.

spin measurements	Fisher inf. incomp. cond.	FT point comp. cond.
coplanar, equiangular vectors	$\eta > \frac{1}{\sqrt{2}} \approx 0.7071$	$\eta < \frac{2}{3} \approx 0.6667$
one vector orth. to the other two vectors	$\eta > 0.72$	$\eta < 0.5907$

Table 1: Comparison between Fisher information incompatibility condition and FT necessary and sufficient compatibility condition

A good perspective about the incompatibility region given by the Fisher information criterion for three unbiased dichotomic qubit measurements can be found using using the generalized Fisher map. It is known that, if the vectors $\vec{n}_k, k = 1, 2, 3$ are mutually orthogonal directions, then the condition (28) reduces to the sufficient and sufficient condition for joint measurability, that is

$$\eta_1^2 + \eta_2^2 + \eta_3^2 \leq 1$$

In the following graphic we aim to analyze the influence of the state ρ on determining the incompatibility region for three unbiased dichotomic qubit measurements, that correspond to σ_x, σ_y and σ_z. This case is presented in Figure 10 using the computation of the height function $\mathfrak{h}(F_\rho(\mathcal{A})) \equiv h(v_1, v_2, v_3, \eta_1, \eta_2, \eta_3)$ which is now a function of the sharpness $\eta_i, i = 1, 2, 3$ and of the parameters of the Bloch vector $v_i, i = 1, 2, 3$.

The simplest case for $\eta_1 = \eta_3 = 0$ that correspond to compatible measurements, for any $\eta_2 \in [0, 1]$, is confirmed by the graphic from Figure 10, as the curves are below the line $y = 2$, which delimits the compatibility and the incompatibility regime. In the setting $\eta_2 = 1/2$ and $\eta_3 = 1/3$, we know that the compatibility condition is $\eta_1 \leq 0.79$, where as for $\eta_2 = \eta_3 = 1/2$, we know that the compatibility condition is $\eta_1 \leq 0.7$. We plot in Figure 10 these two cases (the second and the third graphic), and for different values of the Bloch vector v. As can be seen, in both cases, the incompatibility regime (above $y = 2$) is recovered "faster" for the case of the Bloch vector $v = (0, 0, 0)$, which corresponds to the classical Fisher information map. Moreover, the plots confirm that we cannot detect incompatibility regimes using unit Bloch vector. Indeed, the cyan line that corresponds to the height of the Fisher matrix is always below 2, in this case.

In Figures 11 and 12 we present that the cases of coplanar POVMs, for which we have compared the Zhu incompatibility region and the compatibility region found by FT point. We see that in both cases the best way to detect incompatibility is to use the maximally mixed state, whereas ρ being pure cannot detect incompatibility.

We leave open the existence of a set of measurements which can be shown to be incompatible using the generalized Fisher information map F_ρ for $\rho \neq \mathbb{1}$, but for which the standard map F cannot detect incompatibility. The set of examples we considered in this section show that in most cases of interest, the standard map F performs better that its shifted variant F_ρ.

7 Conclusions and perspectives

The post-processing order of quantum measurements is a natural and important relation, describing intrinsic noise in measurements and giving an operational definition for quantum incompatibility. We have presented a general framework where the set of (equivalence classes of) measurements is mapped into matrices and the post-processing order translates

to the usual matrix ordering. We have illustrated that it depends on the chosen order morphism how much information is lost. Some reasonably looking maps loose too much information and cannot give any criteria for incompatibility, even if they can still detect some non-ordering. In the ideal case, one would get rid of uninteresting details while keeping enough essential information to, e.g., decide incompatibility with high accuracy. The Fisher information map introduced by Zhu is an example of a good and useful order preserving map. It is indeed optimal among quadratic maps as we have shown.

The presented general setting indicates directions for future studies. Firstly, one can seek modifications of the Fisher information map and try to find efficient incompatibility tests by using them. For example, the generalized Fisher information map is such a generalization which could be used to detect incompatibility where the usual map cannot. Instead of using the identity matrix, one could use other quantum states ρ. Full rank states should lead to interesting criteria, since the incertitude about a lower rank state is smaller, and thus the Fisher information matrix should be less informative. Moreover, we have shown that unit rank states are useless for incompatibility detection, so invertible states ρ should be investigated. On the other hand, most quantum measurements of interest have some kind symmetry; the most informative state ρ should also enjoy the same symmetry, making the choice $\rho = \mathbb{1}/d$ reasonable. We leave the question of the hierarchy between the different Fisher information maps F_ρ open for further investigation. It is also possible to investigate order preserving maps that are not quadratic.

Secondly, the underlying idea of order preserving maps that simplify the structure while keeping essential features can be used for other quantum devices than just for quantum measurements. In particular, preparators and instruments have their corresponding post-processing orders and one can formulate analogous maps for them. There, again, the aim would be to forget uninteresting details but keep the essential structure leading to physically relevant phenomena such as incompatibility.

References

[ACHT09] S.T. Ali, C. Carmeli, T. Heinosaari, and A. Toigo. Commutative POVMs and fuzzy observables. *Found. Phys.*, 39:593–612, 2009.

[AK20] N. Andrejic and R. Kunjwal. Joint measurability structures realizable with qubit measurements: Incompatibility via marginal surgery. *Phys. Rev. Research*, 2:043147, 2020.

[BDK+05] F. Buscemi, G.M. D'Ariano, M. Keyl, P. Perinotti, and R.F. Werner. Clean positive operator valued measures. *J. Math. Phys.*, 46:082109, 2005.

[Bha97] R. Bhatia. *Matrix analysis*, volume 169 of *Graduate Texts in Mathematics*. Springer, 1997.

[BHSS13] P. Busch, T. Heinosaari, J. Schultz, and N. Stevens. Comparing the degrees of incompatibility inherent in probabilistic physical theories. *EPL*, 103:10002, 2013.

[BN18] A. Bluhm and I. Nechita. Joint measurability of quantum effects and the matrix diamond. *J. Math. Phys.*, 59:112202, 2018.

[BV04] S. Boyd and L. Vandenberghe. *Convex optimization. Cambridge University Press*, Cambridge, 2004.

[BŻ06]   I. Bengtsson and K. Życzkowski. *Geometry of quantum states. Cambridge University Press*, Cambridge, 2006.

[Bus86]   P. Busch. Unsharp reality and joint measurements for spin observables. *Phys. Rev. D*, 33:2253–2261, 1986.

[CHT12]   C. Carmeli, T. Heinosaari, and A. Toigo. Informationally complete joint measurements on finite quantum systems. *Phys. Rev. A*, 85:012109, 2012.

[CHT19]   C. Carmeli, T. Heinosaari, and T. Toigo. Quantum incompatibility witnesses. *Phys. Rev. Lett.*, 122:130402, 2019.

[DFK19]   S. Designolle, M. Farkas, and J. Kaniewski. Incompatibility robustness of quantum measurements: a unified framework. *New J. Phys.*, 21:113053, 2019.

[GM00]   R.D. Gill and S. Massar. State estimation for large ensembles. *Phys. Rev. A*, 61:042312, 2000.

[GMSG21]   T. Guff, N.A. McMahon, Y.R. Sanders, and A. Gilchrist. A resource theory of quantum measurements. *J. Phys. A: Math. Theor.*, 54:225301, 2021.

[Hei05]   T. Heinonen. Optimal measurements in quantum mechanics. *Phys. Lett. A*, 346:77–86, 2005.

[HHP12]   E. Haapasalo, T. Heinosaari, and J.-P. Pellonpää. Quantum measurements on finite dimensional systems: relabeling and mixing. *Quantum Inf. Process.*, 11:1751–1763, 2012.

[HJN20]   T. Heinosaari, M.A. Jivulescu, and I. Nechita. Random positive operator valued measures. *J. Math. Phys.*, 61:042202, 2020.

[HP17]   E. Haapasalo and J.-P. Pellonpää. Optimal quantum observables. *J. Math. Phys.*, 58:122104, 2017.

[Har13]   A. Harrow. The church of the symmetric subspace. arXiv:1308.6595

[HRS08]   T. Heinosaari, D. Reitzner, and P. Stano. Notes on joint measurability of quantum observables. *Found. Phys.*, 38:1133–1147, 2008.

[Hru16]   P. Hrubeš. On families of anticommuting matrices. *Linear Algebra and its Applications*, 493:494–507, 2016.

[HW94]   P. Hausladen and W.K. Wootters. A 'pretty good' measurement for distinguishing quantum states. *J. Mod. Opt.*, 41:2385–2390, 1994.

[JP07]   A. Jenčová and S. Pulmannová. How sharp are PV measures? *Rep. Math. Phys.*, 59:257–266, 2007.

[JPV08]   A. Jenčová, S. Pulmannová, and E. Vinceková. Sharp and fuzzy observables on effect algebras. *Int. J. Theor. Phys.*, 47:125–148, 2008.

[Kad51]   R.V. Kadison. Order properties of bounded self-adjoint operators. *Proceedings of the American Mathematical Society*, 2:505–510, 1951.

[Kur15a]   Y. Kuramochi. Construction of the least informative observable conserved by a given quantum instrument. *J. Math. Phys.*, 56:092202, 2015.

[Kur15b]   Y. Kuramochi. Minimal sufficient positive-operator valued measure on a separable Hilbert space. *J. Math. Phys.*, 56:102205, 2015.

[Md90]   H. Martens and W.M. de Muynck. Nonideal quantum measurements. *Found. Phys.*, 20:255–281, 1990.

[New32]   M. Newman. Note on an algebraic theorem of Eddington. *J. London Math. Soc.*, 1:93–99, 1932.

[Sco06]   A.J. Scott. Tight informationally complete quantum measurements. *J. Phys. A: Math. Gen.*, 39:13507, 2006.

[SHSL20] P. Skrzypczyk, M.J. Hoban, A. B. Sainz, and N. Linden. The complexity of compatible measurements. *Phys. Rev. Research*, 2:023292, 2020.

[ULMH16] R. Uola, K. Luoma, T. Moroder, and T. Heinosaari. Adaptive strategy for joint measurements. *Phys. Rev. A*, 94:022109, 2016.

[YO13]   S. Yu and C.H. Oh. Quantum contextuality and joint measurement of three observables of a qubit. arXiv: 1312.6470

[ZE11]   H. Zhu and B.-G. Englert. Quantum state tomography with fully symmetric measurements and product measurements. *Phys. Rev. A*, 84:022327, 2011.

[ZHC16]   H. Zhu, M. Hayashi, and L. Chen. Universal steering criteria. *Phys. Rev. Lett.*, 116:070403, 2016.

[Zhu15]   H. Zhu. Information complementarity: a new paradigm for decoding quantum incompatibility. *Sci. Rep.*, 5:14317, 2015.