

# Testing incompatibility of quantum devices with few states

Teiko Heinosaari

*Department of Physics and Astronomy, University of Turku, Finland and  
Quantum algorithms and software, VTT Technical Research Centre of Finland Ltd*

Takayuki Miyadera and Ryo Takakura

*Department of Nuclear Engineering, Kyoto University, 6158540 Kyoto, Japan*

When observations must come from incompatible devices and cannot be produced by compatible devices? This question motivates two integer valued quantifications of incompatibility, called incompatibility dimension and compatibility dimension. The first one quantifies how many states are minimally needed to detect incompatibility if the test states are chosen carefully, whereas the second one quantifies how many states one may have to use if they are randomly chosen. With concrete examples we show that these quantities have unexpected behaviour with respect to noise.

## I. INTRODUCTION

Quantum information processing, including the exciting fields of quantum communication and quantum computation, is ultimately based on the fact that there are new types of resources that can be utilized in carefully designed information processing protocols. The best known feature of quantum information is that quantum systems can be in superposition and entangled states, and these resources lead to applications such as superdense coding and quantum teleportation. While superposition and entanglement are attributes of quantum states, quantum measurements have also features that can power new type of applications. The best known and most studied property is the incompatibility of pairs (or collections) of quantum measurements [1]. It is crucial e.g. in the BB84 quantum key distribution protocol [2] that the used measurements are incompatible.

From the resource perspective, it is important to quantify the incompatibility. There has been several studies on incompatibility robustness, i.e., how incompatibility is affected by noise. This is motivated by the fact that noise is unavoidable in any actual implementation of quantum devices and similar to other quantum properties (e.g. entanglement), large amount of noise destroys incompatibility. Earlier studies have mostly focused in quantifying noise [3] and finding those pairs or collections of measurements that are most robust to certain types of noise [4], or to find conditions under which all incompatibility is completely erased [5]. In this work we introduce quantifications of incompatibility which are motivated by operational aspect of testing whether a collection of devices is incompatible or not. We focus on two integer valued quantifications of incompatibility, called *compatibility dimension* and *incompatibility dimension*. We formulate these concepts for arbitrary collections of devices. Roughly speaking, the first one quantifies how many states we minimally need to use to detect incompatibility if we choose the test states carefully, whereas the second one quantifies how many (affinely independent) states we may have to use if we cannot control

their choice. We study some of the basic properties of these quantifications of incompatibility and we present several examples to demonstrate their behaviour.

We show that, remarkably, even for the standard example of noisy orthogonal qubit observables the incompatibility dimension has a jump in a point where all noise robustness measures are continuous and indicate nothing special to happen. More precisely, the noise parameter has a threshold value where the number of needed test states to reveal incompatibility shifts from 2 to 3. This means that even in this simple class of incompatible pairs of qubit observables there is a *qualitative* difference in the incompatibility of less noisy and more noisy pairs of observables. An interesting additional fact is that the compatibility dimension of these pairs of observables does not depend on the noise parameter.

For simplicity and clarity, we will restrict to finite dimensional Hilbert spaces and observables with finite number of outcomes. Our definitions apply not only to quantum theory but also to any general probabilistic theory (GPT) [6, 7]. However, for the sake of concreteness we keep the discussion in the realm of quantum theory. The main definitions work in any GPT without any changes. We expect that similar findings as the aforementioned result on noisy orthogonal qubit observables can be made in subsequent studies on other collections of devices.

Related studies have been recently reported in [8–11] in the case of quantum observables. We will explain the interconnections of these studies to ours in Section III once the relevant definitions have been introduced.

## II. (IN)COMPATIBILITY ON A SUBSET OF STATES

A quantum observable is mathematically described as a positive operator valued measure (POVM) [12]. A quantum observable with finite number of outcomes is hence a map  $x \mapsto A(x)$  from the outcome set to the set of linear operators on a Hilbert space. We recall that the compatibility of quantum observables  $A_1, \dots, A_n$

with outcome sets  $X_1, \dots, X_n$  means that there exists an observable  $\mathbf{G}$ , called *joint observable*, defined on the product outcome set  $X_1 \times \dots \times X_n$  such that from an outcome  $(x_1, \dots, x_n)$  of  $\mathbf{G}$ , one can infer outcomes for every  $\mathbf{A}_1, \dots, \mathbf{A}_n$  by ignoring the other outcomes. More precisely, the requirement is that

$$\mathbf{A}_1(x_1) = \sum_{x_2, \dots, x_n} \mathbf{G}(x_1, x_2, \dots, x_n) \quad (1)$$

and similarly for all other observables  $\mathbf{A}_2, \dots, \mathbf{A}_n$ . If  $\mathbf{A}_1, \dots, \mathbf{A}_n$  are not compatible, then they are called *incompatible*. This definition applies in all general probabilistic theories and has, in fact, led to inspiring findings on quantum incompatibility compared to incompatibility in other general probabilistic theories [13–15].

**Example 1.** (*Unbiased qubit observables*) We recall a standard example to fix the notation that we will use in later examples. An unbiased qubit observable is a dichotomic observable with outcomes  $\pm$  and determined by a vector  $\mathbf{a} \in \mathbb{R}^3$ ,  $|\mathbf{a}| \leq 1$  via

$$\mathbf{A}^{\mathbf{a}}(\pm) = \frac{1}{2}(\mathbb{1} \pm \mathbf{a} \cdot \boldsymbol{\sigma}),$$

where  $\mathbf{a} \cdot \boldsymbol{\sigma} = a_1\sigma_1 + a_2\sigma_2 + a_3\sigma_3$  and  $\sigma_i$ ,  $i = 1, 2, 3$ , are the Pauli matrices. The norm  $|\mathbf{a}|$  of  $\mathbf{a}$  reflects the noise in  $\mathbf{A}^{\mathbf{a}}$ ; in the extreme case of  $|\mathbf{a}| = 1$  the operators  $\mathbf{A}^{\mathbf{a}}(\pm)$  are projections and the observable is called sharp. As shown in [16], two unbiased qubit observables  $\mathbf{A}^{\mathbf{a}}$  and  $\mathbf{A}^{\mathbf{b}}$  are compatible if and only if

$$|\mathbf{a} + \mathbf{b}| + |\mathbf{a} - \mathbf{b}| \leq 2. \quad (2)$$

There are two extreme cases. Firstly, if  $\mathbf{A}^{\mathbf{a}}$  is sharp then it is compatible with some  $\mathbf{A}^{\mathbf{b}}$  if and only if  $\mathbf{b} = r\mathbf{a}$  for some  $-1 \leq r \leq 1$ . Secondly, if  $|\mathbf{a}| = 0$ , then  $\mathbf{A}^{\mathbf{a}}(\pm) = \frac{1}{2}\mathbb{1}$  and it is called a trivial qubit observable, in which case it is compatible with all other qubit observables.

How can we test if a given family of observables is compatible or incompatible? From the operational point of view, the existence of an observable  $\mathbf{G}$  satisfying (1) is equivalent to the existence of  $\mathbf{G}$  such that for any state  $\varrho$  the equation

$$\text{tr}[\varrho \mathbf{A}_1(x_1)] = \sum_{x_2, \dots, x_n} \text{tr}[\varrho \mathbf{G}(x_1, x_2, \dots, x_n)] \quad (3)$$

holds. To test the incompatibility we should hence check the validity of (3) in a subset of states that spans the whole state space. An obvious question is then if we really need all those states, or if a smaller number of test states is enough. Further, does the number of needed test states depend on the given family of observables? How does noise affect the number of needed test states?

Before contemplating into these questions, we recall that analogous definitions of compatibility and incompatibility make sense for other types of devices, in particular, for instruments and channels [1, 17–22]. We limit

our discussion to quantum devices although, again, the definitions apply to devices in general probabilistic theories. We denote by  $\mathcal{S}(\mathcal{H})$  the set of all density operators on a Hilbert space  $\mathcal{H}$ . The input space of all types of devices must be  $\mathcal{S}(\mathcal{H}_{in})$  on the same Hilbert space  $\mathcal{H}_{in}$  as the devices operate on a same system. We denote  $\mathcal{S}(\mathcal{H}_{in})$  simply by  $\mathcal{S}$ . A device is a completely positive map and the ‘type’ of the device is characterised by its output space. Output spaces for the three basic types of devices are:

- observable:  $P(X) := \{p = \{p(x)\}_{x \in X} \mid 0 \leq p(x) \leq 1, \sum_x p(x) = 1\}$ ,
- channel:  $\mathcal{S}(\mathcal{H}_{out})$ ,
- instrument:  $\mathcal{S}(\mathcal{H}_{out}) \otimes P(X)$ .

We limit our investigation to the cases where the number of outcomes in  $X$  is finite and the output Hilbert space  $\mathcal{H}_{out}$  is finite dimensional.

Two devices  $\mathbf{D}_1$  and  $\mathbf{D}_2$  are *compatible* if there exists a device  $\mathbf{D}$  that can simulate  $\mathbf{D}_1$  and  $\mathbf{D}_2$  simultaneously, meaning that by ignoring disjoint parts of the output of  $\mathbf{D}$  we get the same actions as  $\mathbf{D}_1$  and  $\mathbf{D}_2$ . This kind of device is called a *joint device* of  $\mathbf{D}_1$  and  $\mathbf{D}_2$ . The input space of  $\mathbf{D}$  is the same as for  $\mathbf{D}_1$  and  $\mathbf{D}_2$ , but the output space is the tensor product of their output spaces. If  $\mathbf{D}_1$  and  $\mathbf{D}_2$  are not compatible, then they are *incompatible*. We recall a qubit example to exemplify the general definition.

**Example 2.** (*Unbiased qubit observable and partially depolarizing noise*) A measurement of an unbiased qubit observable  $\mathbf{A}^{\mathbf{a}}$  necessarily disturbs the system. This trade-off is mathematically described by the compatibility relation between observables and channels. Let us consider partially depolarizing qubit channels, which have the form

$$\Gamma_p(\varrho) = p\varrho + (1-p)\frac{1}{2}\mathbb{1} \quad (4)$$

for  $0 \leq p \leq 1$ . A joint device for a channel and observable is an instrument. Hence,  $\mathbf{A}^{\mathbf{a}}$  and  $\Gamma_p$  are compatible if there exists an instrument  $x \mapsto \Phi_x$  such that

$$\sum_x \Phi_x(\varrho) = \Gamma_p(\varrho) \quad \text{and} \quad \text{tr}[\Phi_x(\varrho)] = \text{tr}[\varrho \mathbf{A}^{\mathbf{a}}(x)]$$

for all states  $\varrho$  and outcomes  $x$ . It has been proven in [20] that  $\mathbf{A}^{\mathbf{a}}$  and  $\Gamma_p$  are compatible if and only if

$$|\mathbf{a}| \leq \frac{1}{2} \left( 1 - p + \sqrt{(1-p)(1+3p)} \right). \quad (5)$$

This shows that higher is the norm  $|\mathbf{a}|$ , smaller must  $p$  be.

The earlier discussion motivates the following definition, which is central to our investigation.

**Definition 1.** Let  $\mathcal{S}_0 \subset \mathcal{S}$ . Devices  $D_1, \dots, D_n$  are  $\mathcal{S}_0$ -compatible if there exist compatible devices  $D'_1, \dots, D'_n$  of the same type such that

$$D'_j(\varrho) = D_j(\varrho) \quad (6)$$

for all  $j = 1, \dots, n$  and states  $\varrho \in \mathcal{S}_0$ . Otherwise,  $D_1, \dots, D_n$  are  $\mathcal{S}_0$ -incompatible.

The definition is obviously interesting only when  $D_1, \dots, D_n$  are incompatible in the usual sense, i.e., with respect to the full state space. In that case the definition means that if devices  $D_1, \dots, D_n$  are  $\mathcal{S}_0$ -compatible, their incompatibility cannot be verified by taking test states from  $\mathcal{S}_0$  only, and vice versa, if devices  $D_1, \dots, D_n$  are  $\mathcal{S}_0$ -incompatible, their actions on  $\mathcal{S}_0$  cannot be simulated by any collection of compatible devices and therefore their incompatibility should be able to be observed in some way.

The  $\mathcal{S}_0$ -compatibility depends not only on the size of  $\mathcal{S}_0$  but also on its structure. We start with a simple example showing that there exist sets  $\mathcal{S}_0$  such that an arbitrary family of devices is  $\mathcal{S}_0$ -compatible.

**Example 3.** Any set of devices  $D_1, \dots, D_n$  is  $\mathcal{S}_0$ -compatible if  $\mathcal{S}_0 = \{\varrho_1, \dots, \varrho_k\}$  consists of perfectly distinguishable states. In fact, one may construct a device  $D'_k$  which outputs  $D_k(\varrho_j)$  after confirming an input state is  $\varrho_j$  by measuring an observable that distinguishes the states in  $\mathcal{S}_0$ . It is easy to see that the devices  $D'_1, \dots, D'_n$  are compatible. The same argument works for devices in general probabilistic theories and one can use the same reasoning for a subset  $\mathcal{S}_0$  that is broadcastable [23]. For instance, two qubit states  $\mathbb{1}/2$  and  $|0\rangle\langle 0|$  are broadcastable even though not distinguishable. Any pair of qubit channels  $\Lambda_1$  and  $\Lambda_2$  is  $\mathcal{S}_0$ -compatible for  $\mathcal{S}_0 = \{\mathbb{1}/2, |0\rangle\langle 0|\}$  as we can define  $\Lambda'_j(\varrho) = \sum_{i=0}^1 \langle i|\varrho|i\rangle \Lambda_j(|i\rangle\langle i|)$  for  $j = 1, 2$ . The channel  $\Lambda'_j$  has clearly the same action as  $\Lambda_j$  on  $\mathcal{S}_0$ . A joint channel  $\Lambda$  for  $\Lambda'_1$  and  $\Lambda'_2$  is given as

$$\Lambda(\varrho) = \sum_{i=0}^1 \langle i|\varrho|i\rangle \Lambda_1(|i\rangle\langle i|) \otimes \Lambda_2(|i\rangle\langle i|),$$

and it is clear that, in fact,  $\text{tr}_2[\Lambda(\varrho)] = \Lambda_1(\varrho)$  and  $\text{tr}_1[\Lambda(\varrho)] = \Lambda_2(\varrho)$ .

### III. (IN)COMPATIBILITY DIMENSION OF DEVICES

For a subset  $\mathcal{S}_0 \subset \mathcal{S}$ , we denote by  $\bar{\mathcal{S}}_0$  the intersection of the linear hull of  $\mathcal{S}_0$  with  $\mathcal{S}$ , i.e.,

$$\bar{\mathcal{S}}_0 = \{\varrho \in \mathcal{S} \mid \varrho = \sum_{i=1}^l c_i \varrho_i \text{ for some } c_i \in \mathbb{C} \text{ and } \varrho_i \in \mathcal{S}_0\}$$

In this definition we can assume without restriction that  $c_i \in \mathbb{R}$  and  $\sum_i c_i = 1$  as they follow from the positivity and unit-trace of states. Since the condition (6) is

linear in  $\varrho$ , we conclude that devices  $D_1, \dots, D_n$  are  $\mathcal{S}_0$ -compatible if and only if they are  $\bar{\mathcal{S}}_0$ -compatible. This makes sense: if we can simulate the action of devices for states in  $\mathcal{S}_0$ , we can simply calculate the action for all states that are linear combinations of those states. This observation also shows that a reasonable way to quantify the size of a subset  $\mathcal{S}_0$  for the task in question is the number of affinely independent states.

We consider the following questions. Given a collection of incompatible devices  $D_1, \dots, D_n$ ,

- (a) what is the smallest subset  $\mathcal{S}_0$  such that  $D_1, \dots, D_n$  are  $\mathcal{S}_0$ -incompatible?
- (b) what is the largest subset  $\mathcal{S}_0$  such that  $D_1, \dots, D_n$  are  $\mathcal{S}_0$ -compatible?

Smallest and largest here mean the number of affinely independent states in  $\bar{\mathcal{S}}_0$ . It agrees with the linear dimension of the linear hull of  $\mathcal{S}_0$ , or  $\dim_{\text{aff}} \mathcal{S}_0 + 1$ , where  $\dim_{\text{aff}} \mathcal{S}_0$  is the affine dimension of the affine hull  $\text{aff} \mathcal{S}_0$  of  $\mathcal{S}_0$ . The answer to (a) quantifies how many states we need to use to detect incompatibility if we choose them carefully, whereas the answer to (b) quantifies how many (affinely independent) states we may have to use if we cannot control their choice. Hence for both of these quantities lower number means more incompatibility in the sense of easier detection. The precise mathematical definitions read as follows.

**Definition 2.** For a collection of incompatible devices  $D_1, \dots, D_n$ , we denote

$$\chi_{\text{incomp}}(D_1, \dots, D_n) = \min_{\mathcal{S}_0 \subset \mathcal{S}} \{ \dim_{\text{aff}} \mathcal{S}_0 + 1 \mid D_1, \dots, D_n: \mathcal{S}_0\text{-incompatible} \}$$

and

$$\chi_{\text{comp}}(D_1, \dots, D_n) = \max_{\mathcal{S}_0 \subset \mathcal{S}} \{ \dim_{\text{aff}} \mathcal{S}_0 + 1 \mid D_1, \dots, D_n: \mathcal{S}_0\text{-compatible} \}.$$

We call these numbers the *incompatibility dimension* and *compatibility dimension* of  $D_1, \dots, D_n$ , respectively.

From Example 3 and the fact that the linear dimension of the linear hull of  $\mathcal{S}$  is  $d^2$  we conclude that

$$2 \leq \chi_{\text{incomp}}(D_1, \dots, D_n) \leq d^2 \quad (7)$$

and

$$d \leq \chi_{\text{comp}}(D_1, \dots, D_n) \leq d^2 - 1. \quad (8)$$

Further, from the definitions of these quantities it follows that

$$\chi_{\text{incomp}}(D_1, \dots, D_n) \leq \chi_{\text{comp}}(D_1, \dots, D_n) + 1. \quad (9)$$

We note that based on their definitions, both  $\chi_{\text{incomp}}$  and  $\chi_{\text{comp}}$  are expected to be smaller for collections of devices that are more incompatible.

The following monotonicity property of  $\chi_{\text{incomp}}$  and  $\chi_{\text{comp}}$  under pre-processing is a basic property that any quantification of incompatibility is expected to satisfy.

**Proposition 1.** Let  $\Lambda : \mathcal{S} \rightarrow \mathcal{S}$  be a quantum channel and let  $\tilde{D}_j$  be a pre-processing of  $D_j$  with  $\Lambda$  for each  $j = 1, \dots, n$ , i.e.,  $\tilde{D}_j(\varrho) = D_j(\Lambda(\varrho))$ . If  $\tilde{D}_j$ 's are incompatible, then also  $D_j$ 's are incompatible and

$$\chi_{incomp}(\tilde{D}_1, \dots, \tilde{D}_n) \geq \chi_{incomp}(D_1, \dots, D_n) \quad (10)$$

and

$$\chi_{comp}(\tilde{D}_1, \dots, \tilde{D}_n) \geq \chi_{comp}(D_1, \dots, D_n). \quad (11)$$

*Proof.* Suppose that  $D_1, \dots, D_n$  are  $\mathcal{S}_0$ -compatible for some subset  $\mathcal{S}_0$ . Let  $D$  be a device that gives devices  $\tilde{D}_1, \dots, \tilde{D}_n$  as marginals and these marginals satisfy (6). Then the pre-processing of  $D$  with  $\Lambda$  gives  $\tilde{D}_1, \dots, \tilde{D}_n$  as marginals. The claimed inequalities then follow.  $\square$

The post-processing map of a device  $D$  depends on type of the device. For instance, the output set of an observable is  $P(X)$  and post-processing is then described as a stochastic matrix [24]. We formulate and prove the following monotonicity property of  $\chi_{incomp}$  and  $\chi_{comp}$  under post-processing only for observables. The formulation is analogous for other types of devices.

**Proposition 2.** Let  $\tilde{A}_j$  be a post-processing of  $A_j$  (i.e.  $\tilde{A}_j(x') = \sum_x \nu_j(x', x) A_j(x)$  for some stochastic matrix  $\nu_j$ ) for each  $j = 1, \dots, n$ . If  $\tilde{A}_j$ 's are  $\mathcal{S}_0$ -incompatible, then also  $A_j$ 's are  $\mathcal{S}_0$ -incompatible and

$$\chi_{incomp}(\tilde{A}_1, \dots, \tilde{A}_n) \geq \chi_{incomp}(A_1, \dots, A_n) \quad (12)$$

and

$$\chi_{comp}(\tilde{A}_1, \dots, \tilde{A}_n) \geq \chi_{comp}(A_1, \dots, A_n). \quad (13)$$

*Proof.* Suppose that  $A_1, \dots, A_n$  are  $\mathcal{S}_0$ -compatible for some subset  $\mathcal{S}_0$ . This means that there exists an observable  $G$  satisfying for all  $\varrho \in \mathcal{S}_0$ , any  $j$  and  $x_j$ ,

$$\text{tr}[\varrho A_j(x_j)] = \sum_{l \neq j} \sum_{x_l} \text{tr}[\varrho G(x_1, \dots, x_n)]. \quad (14)$$

We define an observable  $\tilde{G}$  as  $\tilde{G}(x'_1, \dots, x'_n) = \sum_{x_1, \dots, x_n} \nu(x'_1|x_1) \cdots \nu(x'_n|x_n) G(x_1, \dots, x_n)$ , and it then satisfies

$$\text{tr}[\varrho \tilde{A}_j(x'_j)] = \sum_{l \neq j} \sum_{x'_l} \text{tr}[\varrho \tilde{G}(x'_1, \dots, x'_n)] \quad (15)$$

for all  $\varrho \in \mathcal{S}_0$ , any  $j$  and  $x'_j$ . This shows that  $\tilde{A}_1, \dots, \tilde{A}_n$  are  $\mathcal{S}_0$ -compatible. The claimed inequalities then follow.  $\square$

We will now have some examples to demonstrate the values of  $\chi_{incomp}$  and  $\chi_{comp}$  in some standard cases.

**Example 4.** Let us consider the identity channel  $\text{id} : \mathcal{S}(\mathbb{C}^d) \rightarrow \mathcal{S}(\mathbb{C}^d)$ . It follows from the definitions that two identity channels are  $\mathcal{S}_0$ -compatible if and only if  $\mathcal{S}_0$  is a broadcastable set. It is known that a subset of states is broadcastable only if the states commute with each other [25], and for this reason the pair of two identity channels is  $\mathcal{S}_0$ -incompatible whenever  $\mathcal{S}_0$  contains two noncommuting states. Therefore, we have  $\chi_{incomp}(\text{id}, \text{id}) = 2$ . On the other hand,  $\mathcal{S}_0$  consisting of distinguishable states makes the identity channels  $\mathcal{S}_0$ -compatible. As  $\mathcal{S}_0$  consisting of commutative states has at most  $d$  affinely independent states, we conclude that  $\chi_{comp}(\text{id}, \text{id}) = d$ .

A comparison of the results of Example 4 to the bounds (7) and (8) shows that the pair of identity channels has the smallest possible incompatibility and compatibility dimensions. This is quite expectable as that pair is consider to be the most incompatible pair - any device can be post-processed from the identity channel. Perhaps surprisingly, the lower bound of  $\chi_{incomp}$  can be attained already with a pair of dichotomic observables; this is shown in the next example.

**Example 5.** Let  $P$  and  $Q$  be two noncommuting one-dimensional projections in a  $d$ -dimensional Hilbert space  $\mathcal{H}$ . We define two dichotomic observables  $A$  and  $B$  as

$$A(1) = P, A(0) = \mathbb{1} - P, \quad B(1) = Q, B(0) = \mathbb{1} - Q.$$

Let us then consider a subset consisting of two states,

$$\mathcal{S}_0 = \{\varrho^P, \varrho^Q\} := \left\{ \frac{1}{d-1}(\mathbb{1} - P), \frac{1}{d-1}(\mathbb{1} - Q) \right\}.$$

We find that the dichotomic observables  $A$  and  $B$  are  $\mathcal{S}_0$ -incompatible. To see this, let us make a counter assumption that  $A$  and  $B$  are  $\mathcal{S}_0$ -compatible, in which case there exists  $G$  such that the marginal condition (3) holds for both observables and for all  $\varrho \in \mathcal{S}_0$ . We have  $\text{tr}[\varrho^P A(1)] = 0$  and therefore

$$0 = \text{tr}[(\mathbb{1} - P)G(1, 1)] = \text{tr}[(\mathbb{1} - P)G(1, 0)].$$

It follows that  $G(1, 1) = \alpha P$  and  $G(1, 0) = \beta P$ . Further,  $\text{tr}[PA(1)] = 1$  and hence  $\alpha + \beta = 1$ . In a similar way we obtain  $G(1, 1) = \gamma Q$  and  $G(0, 1) = \delta Q$  with  $\gamma + \delta = 1$ . It follows that  $\alpha = \gamma = 0$  and  $\beta = \delta = 1$ . But  $G(1, 0) + G(0, 1) = P + Q$  contradicts  $G(1, 0) + G(0, 1) \leq \mathbb{1}$ . Thus we conclude  $\chi_{incomp}(A, B) = 2$ .

For two incompatible sharp qubit observables (Example 1) the previous example gives a concrete subset of two states such that the observables are incompatible and proves that  $\chi_{incomp}(A^a, A^b) = 2$  for such a pair. The incompatibility dimension for unsharp qubit observables is more complicated and will be treated in Sec. V.

**Example 6.** Let us consider two observables  $A$  and  $B$ . Fix a state  $\varrho_0 \in \mathcal{S}$  and define

$$\mathcal{S}_0 = \{\varrho \in \mathcal{S} : \text{tr}[\varrho A(x)] = \text{tr}[\varrho_0 A(x)] \quad \forall x\}.$$

Then  $A$  and  $B$  are  $\mathcal{S}_0$ -compatible. To see this, we define an observable  $G$  as

$$G(x, y) = \text{tr} [\varrho_0 A(x)] B(y).$$

It is then straightforward to verify that (3) holds for all  $\varrho \in \mathcal{S}_0$ .

As a special instance of this construction, let  $A^{\mathbf{a}}$  be a qubit observable and  $\mathbf{a} \neq 0$  (see Example 1). We choose  $\mathcal{S}_0 = \{\varrho \in \mathcal{S} \mid \text{tr} [\varrho A^{\mathbf{a}}(+)] = \frac{1}{2}\}$ . We then have  $\mathcal{S}_0 = \{\frac{1}{2}(\mathbb{1} + \mathbf{r} \cdot \sigma) \mid \mathbf{r} \cdot \mathbf{a} = 0\}$  and hence  $\dim \text{aff} \mathcal{S}_0 = 2$ . Based on the previous argument,  $A^{\mathbf{a}}$  is  $\mathcal{S}_0$ -compatible with any  $A^{\mathbf{b}}$ . Therefore,  $\chi_{\text{comp}}(A^{\mathbf{a}}, A^{\mathbf{b}}) = 3$  for all incompatible qubit observables  $A^{\mathbf{a}}$  and  $A^{\mathbf{b}}$ .

### Remark on other formulations of incompatibility dimension

The notion of  $\mathcal{S}_0$ -compatibility for quantum observables has been introduced in [8] and in that particular case (i.e. quantum observables) it is equivalent to Def. 1. In the current investigation our focus is on the largest or smallest  $\mathcal{S}_0$  on which devices  $D_1, \dots, D_n$  are compatible or incompatible, and this has some differences to the earlier approaches. In [10], the term ‘‘compatibility dimension’’ was introduced and for observables  $A_1, \dots, A_n$  on a  $d$  dimensional Hilbert space  $\mathcal{H} = \mathbb{C}^d$  it is

$$R(A_1, \dots, A_n) = \max\{r \leq d \mid \exists V: \mathbb{C}^r \rightarrow \mathbb{C}^d \text{ isometry} \\ \text{s.t. } V^* A_1 V, \dots, V^* A_n V \text{ are compatible}\},$$

Evaluations of  $R(A_1, \dots, A_n)$  in various cases such as  $n = 2$  and  $A_1$  and  $A_2$  are rank-1 were presented in [10]. To describe it in our notions, let us denote  $\mathbb{C}^r$  by  $\mathcal{K}$ , and define  $\mathcal{S}_{\mathcal{H}}$  and  $\mathcal{S}_{\mathcal{K}}$  as the set of all density operator on  $\mathcal{H}$  and  $\mathcal{K}$  respectively. We also introduce  $\mathcal{S}_{V\mathcal{K}}$  as

$$\mathcal{S}_{V\mathcal{K}} := \{\rho \in \mathcal{S} \mid \text{supp} \rho \subset V\mathcal{K}\} = V\mathcal{S}_{\mathcal{K}}V^* \subset \mathcal{S}_{\mathcal{H}}.$$

Then, we can see that the  $\mathcal{S}_{\mathcal{K}}$ -compatibility of  $V^* A_1 V, \dots, V^* A_n V$  is equivalent to the  $\mathcal{S}_{V\mathcal{K}}$ -compatibility of  $A_1, \dots, A_n$ . Therefore, if we focus only on sets of states such as  $\mathcal{S}_{V\mathcal{K}}$  (i.e. states with fixed support), then there is no essential difference between our compatibility dimension and the previous one:  $R(A_1, \dots, A_n) = r$  iff  $\chi_{\text{comp}}(A_1, \dots, A_n) = r^2$ . In [10] also the concept of ‘‘strong compatibility dimension’’ was defined as

$$\bar{R}(A_1, \dots, A_n) = \max\{r \leq d \mid \forall V: \mathbb{C}^r \rightarrow \mathbb{C}^d \text{ isometry} \\ \text{s.t. } V^* A_1 V, \dots, V^* A_n V \text{ are compatible}\}.$$

It is related to our notion of incompatibility dimension. In fact, if we only admit sets of states such as  $\mathcal{S}_{V\mathcal{K}}$ , then  $\bar{R}(A_1, \dots, A_n)$  and  $\chi_{\text{incomp}}(A_1, \dots, A_n)$  are essentially the same:  $\bar{R}(A_1, \dots, A_n) = r$  iff  $\chi_{\text{incomp}}(A_1, \dots, A_n) = (r + 1)^2$ .

Similar notions have been introduced and investigated also in [9, 11]. As in [10], these works focus on quantum observables and on subsets of states that are lower dimensional subspaces of the original state space. Therefore, the notions are not directly applicable in GPTs. In [11] incompatibility is classified into three types. They are explained exactly in terms of [10] notion as

- (i) incompressive incompatibility:  $(A_1, \dots, A_n)$  are  $\mathcal{S}_{V\mathcal{K}}$ -compatible for all  $\mathcal{K}$  and  $V$
- (ii) fully compressive incompatibility:  $(A_1, \dots, A_n)$  are  $\mathcal{S}_{V\mathcal{K}}$ -incompatible for all nontrivial  $\mathcal{K}$  and  $V$
- (iii) partly compressive incompatibility: there is a  $V$  and  $\mathcal{K}$  such that  $(A_1, \dots, A_n)$  are  $\mathcal{S}_{V\mathcal{K}}$ -compatible, and some  $V'$  and  $\mathcal{K}'$  such that  $(A_1, \dots, A_n)$  are  $\mathcal{S}_{V'\mathcal{K}'}$ -incompatible. In [11] concrete constructions of these three types of incompatible observables were given.

## IV. AN UPPER BOUND ON THE INCOMPATIBILITY DIMENSION OF OBSERVABLES

An *incompatibility witness* is an affine functional  $\xi$  defined on  $n$ -tuples of observables such that  $\xi$  takes non-negative values on all compatible  $n$ -tuples and a negative value at least for some incompatible  $n$ -tuple [26–28]. Every incompatibility witness can be written in the form

$$\xi(A_1, \dots, A_n) = \delta - \sum_{j=1}^n \sum_{x_j=1}^{m_j} p_j(x_j) \text{tr} [\varrho_{j,x_j} A_j(x_j)], \quad (16)$$

where  $\delta \in \mathbb{R}$  is a constant,  $p_1, \dots, p_n$  are probability distributions,  $\varrho_{j,x_j}$  are states, and  $m_j$  is the number of outcomes of  $A_j$ . This result has been proven in [27] for incompatibility witnesses acting on pairs of observables and the generalization to  $n$ -tuples is straightforward. A witness  $\xi$  *detects* the incompatibility of observables  $A_1, \dots, A_n$  if  $\xi(A_1, \dots, A_n) < 0$ . Assume that  $\xi$  has the form (16) and it detects the incompatibility of observables  $A_1, \dots, A_n$ . Then we can conclude that  $A_1, \dots, A_n$  are  $\mathcal{S}_0$ -incompatible for  $\mathcal{S}_0 = \{\varrho_{j,x_j} \mid j = 1, \dots, n, x_j = 1, \dots, m_j\}$ . Namely, if  $A_1, \dots, A_n$  would be  $\mathcal{S}_0$ -compatible, then we would have compatible observables  $\tilde{A}_1, \dots, \tilde{A}_n$  such that  $\text{tr} [\varrho A_j(x_j)] = \text{tr} [\varrho \tilde{A}_j(x_j)]$  for all  $\varrho \in \mathcal{S}_0$ . This would imply that

$$\xi(A_1, \dots, A_n) = \xi(\tilde{A}_1, \dots, \tilde{A}_n) \geq 0,$$

which contradicts the assumption that  $\xi$  detects the incompatibility of observables  $A_1, \dots, A_n$ . It has been shown in [27] that any incompatible pair of observables is detected by some incompatibility witness. The proof is again straightforward to generalize to  $n$ -tuples of observables and we therefore conclude that

$$\chi_{\text{incomp}}(A_1, \dots, A_n) \leq m_1 + \dots + m_n.$$

We give a better upper bound for the incompatibility dimension by slightly modifying the previous argument that

uses incompatibility witness. The functional  $\xi$  in (16) is of the form

$$\xi(\oplus_j \mathbf{A}_j) = \delta - f(\oplus_j \mathbf{A}_j),$$

with  $\delta \in \mathbb{R}$  and a linear functional  $f$  on  $\oplus_{j=1}^n \mathcal{L}_s(\mathcal{H})^{m_j}$ , where  $\mathcal{L}_s(\mathcal{H})$  is the set of all self-adjoint operators on  $\mathcal{H}$ .

Let us study possible concrete forms of  $f$ . Riesz representation theorem shows that it can be represented as  $f(\oplus_j \mathbf{A}_j) = \sum_{j=1}^n \sum_{x_j}^{m_j} \langle F_j(x_j) | \mathbf{A}_j(x_j) \rangle_{HS}$  with some  $F_j(x_j) \in \mathcal{L}_s(\mathcal{H})$  ( $j = 1, \dots, n$ ,  $x_j = 1, \dots, m_j$ ), where  $\langle F_j(x_j) | \mathbf{A}_j(x_j) \rangle_{HS} := \text{tr}[F_j(x_j) \mathbf{A}_j(x_j)]$  is the Hilbert-Schmidt inner product.

If we define  $F'_j(x_j) = F_j(x_j) + \epsilon_j \mathbb{1}$ , we find

$$\xi(\oplus_j \mathbf{A}_j) = \delta + d \sum_j \epsilon_j - \sum_{j=1}^n \sum_{x_j=1}^{m_j} \langle F'_j(x_j) | \mathbf{A}_j(x_j) \rangle_{HS}.$$

We choose  $\epsilon_j$  so that

$$\sum_{x_j} \text{tr}[F'_j(x_j)] = \sum_{x_j} \langle F'_j(x_j) | \mathbb{1} \rangle_{HS} = 0$$

holds. The choice of  $\{F'_j(x_j)\}_{x_j, j}$  has still some freedom. Each  $F'_j(x_j)$  can be replaced with  $F''_j(x_j) = F'_j(x_j) + T_j$ , where  $T_j \in \mathcal{L}_s(\mathcal{H})$  satisfies  $\text{tr}[T_j] = \langle T_j | \mathbb{1} \rangle_{HS} = 0$ . In fact, it holds that

$$\begin{aligned} & \sum_{x_j} \langle F''_j(x_j) | \mathbf{A}_j(x_j) \rangle_{HS} \\ &= \sum_{x_j} \langle F'_j(x_j) | \mathbf{A}_j(x_j) \rangle_{HS} + \sum_{x_j} \langle T_j | \mathbf{A}_j(x_j) \rangle_{HS} \\ &= \sum_{x_j} \langle F'_j(x_j) | \mathbf{A}_j(x_j) \rangle_{HS} + \langle T_j | \mathbb{1} \rangle_{HS} \\ &= \sum_{x_j} \langle F'_j(x_j) | \mathbf{A}_j(x_j) \rangle_{HS}. \end{aligned}$$

We choose  $T_j$  as  $m_j T_j = -\sum_{x_j=1}^{m_j} F'_j(x_j)$  which indeed satisfies  $m_j \langle T_j | \mathbb{1} \rangle_{HS} = -\sum_{x_j=1}^{m_j} \langle F'_j(x_j) | \mathbb{1} \rangle_{HS} = 0$ , i.e.  $\text{tr}[T_j] = 0$ , to obtain

$$\sum_{x_j} F''_j(x_j) = O.$$

We further choose large numbers  $\alpha_j \geq 0$  so that  $G_j(x_j) := F''_j(x_j) + \alpha_j \mathbb{1} \geq O$  for all  $j$  and  $x_j$ . Now we obtain a representation of the witness which is equivalent to  $\xi$  for  $n$ -tuples of observables as

$$\begin{aligned} \xi^*(\oplus_j \mathbf{A}_j) &= \delta + d \sum_j (\epsilon_j + \alpha_j) \\ &\quad - \sum_j \sum_{x_j} \langle G_j(x_j) | \mathbf{A}_j(x_j) \rangle_{HS}, \end{aligned}$$

where positive operators  $G_j(x_j)$ 's satisfy  $\sum_{x_j} G_j(x_j) = m_j \alpha_j \mathbb{1}$ . Defining density operators  $\rho_j(x_j)$  by  $\rho_j(x_j) =$

$\frac{G_j(x_j)}{\text{tr}[G_j(x_j)]}$ , we obtain yet another representation

$$\begin{aligned} \xi^*(\oplus_j \mathbf{A}_j) &= \delta + d \sum_j (\epsilon_j + \alpha_j) \\ &\quad - \sum_j \sum_{x_j} \text{tr}[G_j(x_j)] \text{tr}[\rho_j(x_j) \mathbf{A}_j(x_j)] \end{aligned}$$

with  $\rho_j(x_j)$ 's satisfying constraints

$$\sum_{x_j} \text{tr}[G_j(x_j)] \rho_j(x_j) = m_j \alpha_j \mathbb{1}. \quad (17)$$

Thus, similarly to the arguments in the beginning of this section,  $\mathbf{A}_1, \dots, \mathbf{A}_n$  are  $\mathcal{S}_0$ -incompatible with  $\mathcal{S}_0 = \{\rho_j(x_j)\}_{j, x_j}$ . To evaluate  $\text{dim aff } \mathcal{S}_0$ , we focus on the condition (17). Introducing parameters  $p_j(x_j) := \text{tr}[G_j(x_j)]/d m_j \alpha_j$  such that  $\sum_{x_j} p_j(x_j) = 1$ , we obtain

$$\sum_{x_j} p_j(x_j) \rho_j(x_j) = \frac{1}{d} \mathbb{1},$$

or

$$\sum_{x_j} p_j(x_j) \tilde{\rho}_j(x_j) = 0,$$

where  $\tilde{\rho}_j(x_j) := \rho_j(x_j) - \frac{1}{d} \mathbb{1}$ . It follows that  $\{\tilde{\rho}_j(x_j)\}_{x_j}$  are linearly dependent, and thus

$$\text{dimspan}\{\tilde{\rho}_j(x_j)\}_{x_j} \leq m_1 - 1.$$

Similar arguments for the other  $j$ 's result in

$$\text{dimspan}\{\tilde{\rho}_j(x_j)\}_{j, x_j} \leq \sum_j (m_j - 1) = \sum_j m_j - n.$$

Considering that

$$\text{dimspan}\{\tilde{\rho}_j(x_j)\}_{j, x_j} = \text{dim aff}\{\rho_j(x_j)\}_{j, x_j}$$

holds, we can obtain the following claim.

**Proposition 3.** *Let  $\mathbf{A}_1, \dots, \mathbf{A}_n$  be incompatible observables with  $m_1, \dots, m_n$  outcomes, respectively. Then*

$$\chi_{\text{incomp}}(\mathbf{A}_1, \dots, \mathbf{A}_n) \leq \sum_{j=1}^n m_j - n + 1$$

The bound is not tight in general since the right-hand side of the above inequality can exceed the bound obtained in (7). However, for small  $n$  and  $m_j$ 's, the bound can be tight. In fact, while for  $n = 2$  and  $m_1 = m_2 = 2$  it gives  $\chi_{\text{incomp}}(\mathbf{A}_1, \mathbf{A}_2) \leq 3$ , we will construct an example which attains this upper bound in the next section.

## V. MUTUALLY UNBIASED QUBIT OBSERVABLES

In this section we study the incompatibility dimension of pairs of unbiased qubit observables introduced in Example 1. We concentrate on pairs that are mutually unbiased, i.e.,  $\text{tr}[A^a(\pm)A^b(\pm)] = 1/2$ . (This terminology originates from the fact that if the observables are sharp, then the respective orthonormal bases are mutually unbiased. In the previously written form the definition makes sense also for unsharp observables [29].) The condition of mutual unbiasedness is invariant under a global unitary transformation, hence it is enough to fix the basis  $\mathbf{x} = (1, 0, 0)$ ,  $\mathbf{y} = (0, 1, 0)$ ,  $\mathbf{z} = (0, 0, 1)$  in  $\mathbb{R}^3$  and choose two of these unit vectors. We will study the observables  $A^{t\mathbf{x}}$  and  $A^{t\mathbf{y}}$ , where  $0 \leq t \leq 1$ . The observables are written explicitly as

$$A^{t\mathbf{x}}(\pm) = \frac{1}{2}(\mathbb{1} \pm t\sigma_1), \quad A^{t\mathbf{y}}(\pm) = \frac{1}{2}(\mathbb{1} \pm t\sigma_2).$$

The condition (2) shows that  $A^{t\mathbf{x}}$  and  $A^{t\mathbf{y}}$  are incompatible if and only if  $1/\sqrt{2} < t \leq 1$ . The choice of having mutually unbiased observables as well as using a single noise parameter instead of two is to simplify the calculations.

We have seen in Example 6 that  $\chi_{\text{comp}}(A^{t\mathbf{x}}, A^{t\mathbf{y}}) = 3$  for all values  $t$  for which the pair is incompatible. We have further seen (discussion after Example 5) that  $\chi_{\text{incomp}}(A^{\mathbf{x}}, A^{\mathbf{y}}) = 2$ , and from Prop. 3 follows that  $\chi_{\text{incomp}}(A^{t\mathbf{x}}, A^{t\mathbf{y}}) \leq 3$  for all  $1/\sqrt{2} < t \leq 1$ . The remaining question is then about the exact value of  $\chi_{\text{incomp}}(A^{t\mathbf{x}}, A^{t\mathbf{y}})$ , which can depend on the noise parameter  $t$ .

Let us first make a simple observation that follows from Prop. 2. Considering that  $A^{s\mathbf{x}}$  is obtained as a post-processing of  $A^{t\mathbf{x}}$  if and only if  $s \leq t$ , we conclude that

$$\begin{aligned} \chi_{\text{incomp}}(A^{s\mathbf{x}}, A^{s\mathbf{y}}) &= 2 \\ \Rightarrow \chi_{\text{incomp}}(A^{t\mathbf{x}}, A^{t\mathbf{y}}) &= 2 \quad \text{for } s \leq t, \end{aligned}$$

and

$$\begin{aligned} \chi_{\text{incomp}}(A^{s\mathbf{x}}, A^{s\mathbf{y}}) &= 3 \\ \Rightarrow \chi_{\text{incomp}}(A^{t\mathbf{x}}, A^{t\mathbf{y}}) &= 3 \quad \text{for } s \geq t. \end{aligned}$$

Interestingly, there is a threshold value  $t_0$  where the value of  $\chi_{\text{incomp}}(A^{t\mathbf{x}}, A^{t\mathbf{y}})$  changes; this is the content of the following proposition.

**Proposition 4.** *There exists  $1/\sqrt{2} < t_0 < 1$  such that  $\chi_{\text{incomp}}(A^{t\mathbf{x}}, A^{t\mathbf{y}}) = 3$  for  $1/\sqrt{2} < t \leq t_0$  and  $\chi_{\text{incomp}}(A^{t\mathbf{x}}, A^{t\mathbf{y}}) = 2$  for  $t_0 < t \leq 1$ .*

The main line of the lengthy proof of Prop. 4 is the following. Defining two subsets  $L$  and  $M$  of  $(\frac{1}{\sqrt{2}}, 1]$  as

$$\begin{aligned} L &:= \{t \mid \chi_{\text{incomp}}(A^{t\mathbf{x}}, A^{t\mathbf{y}}) = 2\}, \\ M &:= \{t \mid \chi_{\text{incomp}}(A^{t\mathbf{x}}, A^{t\mathbf{y}}) = 3\}, \end{aligned} \quad (18)$$

we see that

$$\inf L = \sup M (=: t'_0) \quad (19)$$

holds unless  $L$  and  $M$  are empty. By its definition, the number  $t'_0$  satisfies

$$\begin{aligned} \chi_{\text{incomp}}(A^{t\mathbf{x}}, A^{t\mathbf{y}}) &= 2 \quad \text{for } t > t'_0, \\ \chi_{\text{incomp}}(A^{t\mathbf{x}}, A^{t\mathbf{y}}) &= 3 \quad \text{for } t < t'_0. \end{aligned}$$

Based on the considerations above, the proof of Proposition 4 proceeds as follows. First, in Part 1, we prove that  $M$  is nonempty while  $L$  has already been shown to be nonempty as  $t = 1 \in L$ . It will be found that  $\chi_{\text{incomp}}(A^{t\mathbf{x}}, A^{t\mathbf{y}}) = 3$  for  $t$  sufficiently close to  $\frac{1}{\sqrt{2}}$ , and thus  $t'_0$  introduced above can be defined successfully. Then, we demonstrate in Part 2 that  $\sup M = \max M$ , i.e.  $t'_0$  equals to  $t_0$  in the claim of Prop. 4.

**Remark 1.** In [8] a similar problem to ours was considered. While in that work the focus was on several affine sets, and a threshold value  $t_0$  was given for each of them by means of their semidefinite programs where observables  $\{A^{t\mathbf{x}}, A^{t\mathbf{y}}, A^{t\mathbf{z}}\}$  become compatible, we are considering *all* affine sets with dimension 2.

### Proof of Proposition 4 : Part 1

In order to prove that  $M$  is nonempty, let us introduce some relevant notions:

$$\begin{aligned} D &:= \{\mathbf{v} \mid |\mathbf{v}| \leq 1, v_z = 0\} \subset B := \{\mathbf{v} \mid |\mathbf{v}| \leq 1\}, \\ \mathcal{S}_D &:= \{\rho^{\mathbf{v}} \mid \mathbf{v} \in D\} \subset \mathcal{S} = \{\rho^{\mathbf{v}} \mid \mathbf{v} \in B\}, \end{aligned}$$

where  $\mathbf{v} = v_x\mathbf{x} + v_y\mathbf{y} + v_z\mathbf{z} \in \mathbb{R}^3$ , and  $\rho^{\mathbf{v}} := \frac{1}{2}(\mathbb{1} + \mathbf{v} \cdot \sigma)$ . Since  $\mathcal{S}_D$  is a convex set, we can treat  $\mathcal{S}_D$  almost like a quantum system. In the following, we will do it without giving precise definitions because they are obvious. For an observable  $E$  on  $\mathcal{S}$  with effects  $\{E(x)\}_x$ , we write its restriction to  $\mathcal{S}_D$  as  $E|_D$  with effects  $\{E(x)|_D\}_x$ , which is an observable on  $\mathcal{S}_D$ . It is easy to obtain the following Lemma.

**Lemma 1.** *The followings are equivalent:*

- (i)  $A^{t\mathbf{x}}$  and  $A^{t\mathbf{y}}$  are incompatible (thus  $\frac{1}{\sqrt{2}} < t \leq 1$ ).
- (ii)  $A^{t\mathbf{x}}$  and  $A^{t\mathbf{y}}$  are  $\mathcal{S}_D$ -incompatible.
- (iii)  $A^{t\mathbf{x}}|_D$  and  $A^{t\mathbf{y}}|_D$  are incompatible as observables on  $\mathcal{S}_D$ .

*Proof.* (i)  $\Rightarrow$  (iii). Suppose that  $A^{t\mathbf{x}}|_D$  and  $A^{t\mathbf{y}}|_D$  are compatible in  $\mathcal{S}_D$ . There exists an observable  $M$  on  $\mathcal{S}_D$  whose marginals coincide with  $A^{t\mathbf{x}}|_D$  and  $A^{t\mathbf{y}}|_D$ . One can extend this  $M$  to the whole  $\mathcal{S}$  so that it does not depend on  $z$  (for example, one can simply regard its effect  $c_0\mathbb{1} + c_1\sigma_1 + c_2\sigma_2$  as an effect on  $\mathcal{S}$ ). Since both  $A^{t\mathbf{x}}|_D$  and  $A^{t\mathbf{y}}|_D$  also do not depend on  $z$ , the extension of  $M$

gives a joint observable of  $A^{tx}$  and  $A^{tx}$ .

(iii)  $\Rightarrow$  (ii). Suppose that  $A^{tx}$  and  $A^{ty}$  are  $\mathcal{S}_D$ -compatible. There exists an observable  $M$  on  $\mathcal{S}$  whose marginals coincide with  $A^{tx}$  and  $A^{ty}$  in  $\mathcal{S}_D$ . The restriction of  $M$  on  $\mathcal{S}_D$  proves that (iii) is false.

(ii)  $\Rightarrow$  (i). Suppose that  $A^{tx}$  and  $A^{ty}$  are compatible, then they are  $\mathcal{S}_D$ -compatible.  $\square$

This lemma demonstrates that the incompatibility of  $A^{tx}$  and  $A^{ty}$  means the incompatibility of  $A^{tx}|_D$  and  $A^{ty}|_D$ . We can present further observations.

**Lemma 2.** *Let us consider two pure states  $\rho^{r_1}$  and  $\rho^{r_2}$  ( $\mathbf{r}_1, \mathbf{r}_2 \in \partial B$ ,  $\mathbf{r}_1 \neq \mathbf{r}_2$ ), and a convex subset  $\mathcal{S}_0$  of  $\mathcal{S}$  generated by them:  $\mathcal{S}_0 := \{p\rho^{r_1} + (1-p)\rho^{r_2} \mid 0 \leq p \leq 1\}$ . We also introduce an affine projection  $P$  by  $P\rho^{\mathbf{v}} = \rho_{P\mathbf{v}}$ , where  $\rho^{\mathbf{v}} \in \mathcal{S}$  with  $\mathbf{v} = v_x\mathbf{x} + v_y\mathbf{y} + v_z\mathbf{z}$  and  $P\mathbf{v} = v_x\mathbf{x} + v_y\mathbf{y}$ , and extend it affinely. The affine hull of  $\mathcal{S}_0$  is projected to  $\mathcal{S}_D$  as*

$$P\mathcal{S}_0 := \{\lambda P\rho^{r_1} + (1-\lambda)P\rho^{r_2} \mid \lambda \in \mathbf{R}\} \cap \mathcal{S}_D. \quad (20)$$

If  $A^{tx}$  and  $A^{ty}$  are  $\mathcal{S}_0$ -incompatible, then their restrictions  $A^{tx}|_D$  and  $A^{ty}|_D$  are  $P\mathcal{S}_0$ -incompatible.

*Proof.* Suppose that  $A^{tx}$  and  $A^{ty}$  are  $\mathcal{S}_0$ -incompatible. It implies  $\text{Pr}_1 \neq \text{Pr}_2$  i.e.  $P\rho^{r_1} \neq P\rho^{r_2}$  (see Example 6), and thus  $P\mathcal{S}_0$  is a segment in  $\mathcal{S}_D$ .

If  $A^{tx}|_D$  and  $A^{ty}|_D$  are  $P\mathcal{S}_0$ -compatible, then there exists a joint observable  $M$  on  $\mathcal{S}_D$  such that its marginals coincide with  $A^{tx}|_D$  and  $A^{ty}|_D$  on  $P\mathcal{S}_0 \subset \mathcal{S}_D$ . This  $M$  can be extended to an observable on  $\mathcal{S}$  so that the extension does not depend on  $z$ . Because

$$\begin{aligned} \text{tr}[A^{tx}(\pm)P\rho^{r_1}] &= \text{tr}[A^{tx}(\pm)\rho^{r_1}], \\ \text{tr}[A^{tx}(\pm)P\rho^{r_2}] &= \text{tr}[A^{tx}(\pm)\rho^{r_2}] \end{aligned}$$

(and their  $\mathbf{y}$ -counterparts) hold due to the independence of  $A^{tx}(\pm)$  from  $\sigma_3$ , the marginals of  $M$  coincide with  $A^{tx}$  and  $A^{ty}$  on  $\mathcal{S}_0$ . It results in the  $\mathcal{S}_0$ -compatibility of  $A^{tx}$  and  $A^{ty}$ , which is a contradiction.  $\square$

It follows from this lemma that  $\chi_{incomp}(A^{tx}|_D, A^{ty}|_D)$  is two when  $\chi_{incomp}(A^{tx}, A^{ty})$  is two, equivalently  $\chi_{incomp}(A^{tx}, A^{ty})$  is three when  $\chi_{incomp}(A^{tx}|_D, A^{ty}|_D)$  is three (remember that  $\chi_{incomp}(A^{tx}, A^{ty}) \leq 3$ ). In fact, the converse also holds.

**Lemma 3.**  *$\chi_{incomp}(A^{tx}|_D, A^{ty}|_D)$  is three when  $\chi_{incomp}(A^{tx}, A^{ty})$  is three.*

*Proof.* Let  $\chi_{incomp}(A^{tx}, A^{ty}) = 3$ . It follows that for any line  $S \subset \mathcal{S}$ ,  $A^{tx}$  and  $A^{ty}$  are  $S$ -compatible. In particular,  $A^{tx}$  and  $A^{ty}$  are  $S'$ -compatible for any line  $S'$  in  $\mathcal{S}_D$ , and thus there is an observable  $M$  such that its marginals coincide with  $A^{tx}$  and  $A^{ty}$  on  $S'$ . It is easy to see that the marginals of  $M|_D$  coincide with  $A^{tx}|_D$  and  $A^{ty}|_D$  on  $S'$ , which results in the  $S'$ -compatibility of  $A^{tx}|_D$  and  $A^{ty}|_D$ . Because  $S'$  is arbitrary, we can conclude  $\chi_{incomp}(A^{tx}|_D, A^{ty}|_D) = 3$ .  $\square$

The lemmas above manifest that if  $A^{tx}$  and  $A^{ty}$  are incompatible, then  $A^{tx}|_D$  and  $A^{ty}|_D$  are also incompatible and

$$\chi_{incomp}(A^{tx}, A^{ty}) = \chi_{incomp}(A^{tx}|_D, A^{ty}|_D).$$

Therefore, in the following, we denote  $A^{tx}|_D$  and  $A^{ty}|_D$  simply by  $A_D^{tx}$  and  $A_D^{ty}$  respectively, and focus on the quantity  $\chi_{incomp}(A_D^{tx}, A_D^{ty})$  instead of  $\chi_{incomp}(A^{tx}, A^{ty})$ . Before proceeding to the next step, let us confirm our strategy of this part. It is composed by further two parts: (a) and (b). In (a), we will consider a line (segment)  $\mathcal{S}_1$  in  $\mathcal{S}_D$ , and consider for  $0 < t < 1$  all pairs of observables  $(\tilde{A}_1^t, \tilde{A}_2^t)$  on  $\mathcal{S}_D$  which coincide with  $(A_D^{tx}, A_D^{ty})$  on  $\mathcal{S}_1$ . Then, in (b), we will investigate the (in)compatibility of those  $\tilde{A}_1^t$  and  $\tilde{A}_2^t$  in order to obtain  $\chi_{incomp}(A_D^{tx}, A_D^{ty})$ . It will be shown that when  $t$  is sufficiently small, there exists a compatible pair  $(\tilde{A}_1^t, \tilde{A}_2^t)$  for any  $\mathcal{S}_1$ , that is,  $A_D^{tx}$  and  $A_D^{ty}$  are  $\mathcal{S}_1$ -compatible for any line  $\mathcal{S}_1$ . It results in  $\chi_{incomp}(A_D^{tx}, A_D^{ty}) = 3$ , and thus  $M \neq \emptyset$ .

(a) Let us consider two pure states  $\rho^{r_1}$  and  $\rho^{r_2}$  with  $\mathbf{r}_1, \mathbf{r}_2 \in \partial D$  ( $\mathbf{r}_1 \neq \mathbf{r}_2$ ), and a convex set  $\mathcal{S}_1 := \{p\rho^{r_1} + (1-p)\rho^{r_2} \mid 0 \leq p \leq 1\}$ . We set parameters  $\varphi_1$  and  $\varphi_2$  as

$$\mathbf{r}_1 = \cos \varphi_1 \mathbf{x} + \sin \varphi_1 \mathbf{y}, \quad (21)$$

$$\mathbf{r}_2 = \cos \varphi_2 \mathbf{x} + \sin \varphi_2 \mathbf{y}, \quad (22)$$

where  $-\pi \leq \varphi_1 < \varphi_2 < \pi$ . By exchanging  $\pm$  properly, without loss of generality we can assume the line connecting  $\mathbf{r}_1$  and  $\mathbf{r}_2$  passes through above the origin (instead of below). In this case, from geometric consideration, we have

$$\begin{aligned} 0 < \varphi_2 - \varphi_1 &\leq \pi, \\ 0 &\leq \frac{\varphi_1 + \varphi_2}{2} \leq \frac{\pi}{2}. \end{aligned} \quad (23)$$

Note that when  $\varphi_2 - \varphi_1 = \pi$ ,  $\rho^{r_1}$  and  $\rho^{r_2}$  are perfectly distinguishable, which results in the  $\mathcal{S}_1$ -compatibility of  $A_D^{tx}$  and  $A_D^{ty}$  (see Example 3). On the other hand, when  $\frac{\varphi_1 + \varphi_2}{2} = 0$  or  $\frac{\pi}{2}$ ,  $\text{tr}[\rho A_D^{tx}(+)]$  or  $\text{tr}[\rho A_D^{ty}(+)]$  is constant for  $\rho \in \mathcal{S}_1$  respectively, so  $A_D^{tx}$  and  $A_D^{ty}$  are  $\mathcal{S}_1$ -compatible (see Example 6). Thus, instead of (23), we hereafter assume

$$\begin{aligned} 0 < \frac{\varphi_2 - \varphi_1}{2} &< \frac{\pi}{2}, \\ 0 < \frac{\varphi_1 + \varphi_2}{2} &< \frac{\pi}{2}. \end{aligned} \quad (24)$$

Next, we consider a binary observable  $\tilde{A}_1^t$  on  $\mathcal{S}_D$  which coincides with  $A_D^{tx}$  on  $\mathcal{S}_1 \subset \mathcal{S}_D$ . There are many possible  $\tilde{A}_1^t$ , and each  $\tilde{A}_1^t$  is determined completely by its effect  $\tilde{A}_1^t(+)$  corresponding to the outcome '+' because it is binary. The effect  $\tilde{A}_1^t(+)$  is associated with a vector  $\mathbf{v}_1 \in D$  defined as

$$\mathbf{v}_1 := \text{argmax}_{\mathbf{v} \in D} \text{tr}[\rho_{\mathbf{v}} \tilde{A}_1^t(+)]. \quad (25)$$



Let us introduce a parameter  $\xi_1 \in [-\pi, \pi)$  by

$$\mathbf{v}_1 = \cos \xi_1 \mathbf{x} + \sin \xi_1 \mathbf{y}, \quad (26)$$

and express  $\tilde{\mathbf{A}}_1^t(+)$  as

$$\tilde{\mathbf{A}}_1^t(+)=\frac{1}{2}((1+w(\xi_1))\mathbb{1}+\mathbf{m}_1(\xi_1)\cdot\sigma), \quad (27)$$

where we set

$$\mathbf{m}_1(\xi_1)=C_1(\xi_1)\mathbf{v}_1 \quad \text{with} \quad 0\leq C_1(\xi_1)\leq 1. \quad (28)$$

Because

$$\begin{aligned} \text{tr}[\rho^{\mathbf{r}_1}\mathbf{A}_D^{t\mathbf{x}}(+)] &= \text{tr}[\rho^{\mathbf{r}_1}\tilde{\mathbf{A}}_1^t(+)], \\ \text{tr}[\rho^{\mathbf{r}_2}\mathbf{A}_D^{t\mathbf{y}}(+)] &= \text{tr}[\rho^{\mathbf{r}_2}\tilde{\mathbf{A}}_1^t(+)], \end{aligned}$$

namely

$$\begin{aligned} \frac{1}{2}+\frac{t}{2}\cos\varphi_1 &= \frac{1+w_1(\xi_1)}{2}+\frac{C_1(\xi_1)}{2}\cos(\varphi_1-\xi_1), \\ \frac{1}{2}+\frac{t}{2}\cos\varphi_2 &= \frac{1+w_1(\xi_1)}{2}+\frac{C_1(\xi_1)}{2}\cos(\varphi_2-\xi_1), \end{aligned} \quad (29)$$

hold, we can obtain

$$\begin{aligned} C_1(\xi_1) &= \frac{t(\cos\varphi_1-\cos\varphi_2)}{\cos(\varphi_1-\xi_1)-\cos(\varphi_2-\xi_1)} \\ &= \frac{t\sin\varphi_0}{\sin(\varphi_0-\xi_1)}, \\ w_1(\xi_1) &= -t\left(\frac{\sin(\varphi_1-\varphi_2)}{2\sin(\frac{\varphi_1-\varphi_2}{2})}\right)\times\left(\frac{\sin\xi_1}{\sin(\varphi_0-\xi_1)}\right) \\ &= \frac{-t\cos\psi_0\sin\xi_1}{\sin(\varphi_0-\xi_1)}, \end{aligned} \quad (30)$$

where we set  $\varphi_0 := \frac{\varphi_1+\varphi_2}{2}$  and  $\psi_0 := \frac{\varphi_2-\varphi_1}{2}$  ( $0 < \varphi_0 < \frac{\pi}{2}$ ,  $0 < \psi_0 < \frac{\pi}{2}$ ). Note that if  $\sin(\varphi_0 - \xi_1) = 0$  or  $\cos(\varphi_1 - \xi_1) - \cos(\varphi_2 - \xi_1) = 0$  holds, then  $\cos\varphi_1 - \cos\varphi_2 = 0$  holds (see (30)). It means  $\varphi_0 = 0$ , which is a contradiction, and thus  $\sin(\varphi_0 - \xi_1) \neq 0$  (that is,  $C_1(\xi_1)$  and  $w_1(\xi_1)$  in (30), (31) are well-defined). Moreover, because  $C_1(\xi_1) \geq 0$ , we can see from (30) that  $\sin(\varphi_0 - \xi_1) > 0$  holds, which results in

$$0 \leq \xi_1 < \varphi_0, \quad (32)$$

or

$$-\pi + \varphi_0 < \xi_1 \leq 0. \quad (33)$$

In addition,  $\xi_1$  is restricted also by the condition that  $\tilde{\mathbf{A}}_1^t(\pm)$  are positive. Since the eigenvalues of  $\tilde{\mathbf{A}}_1^t(\pm)$  are  $\frac{1}{2}((1+w_1(\xi_1)) \pm C_1(\xi_1))$ , the restriction comes from both

$$\begin{aligned} 1+w_1(\xi_1)+C_1(\xi_1) &\leq 2, \\ 1+w_1(\xi_1)-C_1(\xi_1) &\geq 0, \end{aligned} \quad (34)$$

equivalently

$$1-w_1(\xi_1)\geq C_1(\xi_1), \quad (35)$$

$$1+w_1(\xi_1)\geq C_1(\xi_1). \quad (36)$$

When (33) (i.e.  $\sin \xi_1 \leq 0$ ) holds,  $w_1(\xi_1) \geq 0$  holds, and thus (35) is sufficient. It is written explicitly as

$$\sin(\varphi_0 - \xi_1) + t \sin \xi_1 \cos \psi_0 \geq t \sin \varphi_0,$$

or

$$\frac{1}{t}\cos\xi_1+\frac{1}{t\sin\varphi_0}(t\cos\psi_0-\cos\varphi_0)\sin\xi_1\geq 1. \quad (37)$$

In order to investigate (37), we adopt a geometric method here while it can be solved in an analytic way. Let us define

$$h_1(t, \varphi_0, \psi_0) = \frac{1}{t \sin \varphi_0} (t \cos \psi_0 - \cos \varphi_0). \quad (38)$$

Then, we can rewrite (37) as

$$(\cos \xi_1, \sin \xi_1) \cdot \left[ \left( \frac{1}{t}, h_1 \right) - (\cos \xi_1, \sin \xi_1) \right] \geq 0. \quad (39)$$

In fact, it can be verified easily that  $(\frac{1}{t}, h_1)$  is the intersection of the line  $l_1 := \{\lambda \mathbf{r}_1 + (1-\lambda)\mathbf{r}_2 \mid \lambda \in \mathbb{R}\}$  and the line  $x = \frac{1}{t}$  in  $\mathbb{R}^2$ . Considering this fact, we can find that  $\xi_1$  satisfies (39) if and only if

$$\xi_1^{\min}(t, \varphi_0, \psi_0) \leq \xi_1 \leq 0, \quad (40)$$

where  $\xi_1^{\min}(t, \varphi_0, \psi_0)$  is determined by the condition

$$\left[ \left( \frac{1}{t}, h_1 \right) - (\cos \xi_1^{\min}, \sin \xi_1^{\min}) \right] \perp (\cos \xi_1^{\min}, \sin \xi_1^{\min}) \quad (41)$$

(see FIG. 1). Analytically, it corresponds to the case when the equality of (37) holds:

$$\frac{1}{t}\cos\xi_1^{\min}+\frac{1}{t\sin\varphi_0}(t\cos\psi_0-\cos\varphi_0)\sin\xi_1^{\min}=1, \quad (42)$$

or

$$1-w_1(\xi_1^{\min})=C_1(\xi_1^{\min}).$$

It can be represented explicitly as

$$\begin{aligned} (t^2\cos^2\psi_0-2t\cos\varphi_0\cos\psi_0+1)\sin^2\xi_1^{\min} \\ -2t\sin\varphi_0(t\cos\psi_0-\cos\varphi_0)\sin\xi_1^{\min} \\ +(t^2-1)\sin^2\varphi_0=0, \end{aligned} \quad (43)$$

and  $\sin \xi_1^{\min}$  is obtained as its negative solution. Note that since the coefficient  $(t^2 \cos^2 \psi_0 - 2t \cos \varphi_0 \cos \psi_0 + 1)$  is strictly positive, the solutions do not show any singular behavior. In summary, we have obtained

$$\xi_1^{\min}(t, \varphi_0, \psi_0) \leq \xi_1 \leq 0 \quad (44)$$

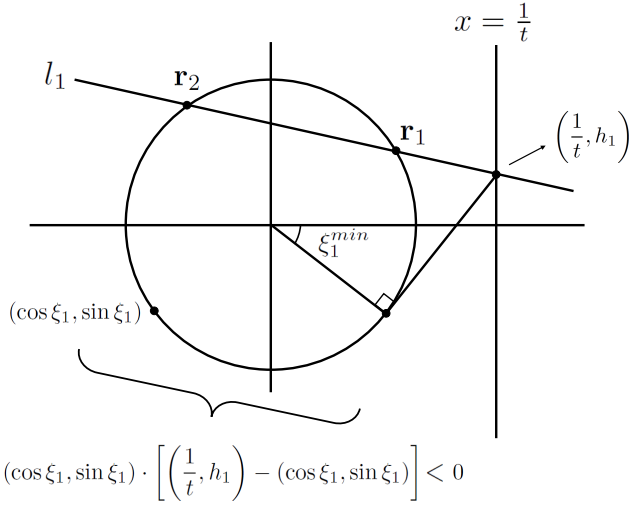


FIG. 1: Geometric description of determining  $\xi_1^{min}$ .

with  $\xi_1^{min}(t, \varphi_0, \psi_0)$  uniquely determined for  $t$ ,  $\varphi_0$ , and  $\psi_0$  by

$$\begin{cases} -\pi + \varphi_0 < \xi_1^{min}(t, \varphi_0, \psi_0) \leq 0, \\ 1 - w_1(\xi_1^{min}(t, \varphi_0, \psi_0)) = C_1(\xi_1^{min}(t, \varphi_0, \psi_0)). \end{cases} \quad (45)$$

On the other hand, when (32) (i.e.  $\sin \xi_1 \geq 0$ ) holds, (36) is sufficient. It results in a tight condition for  $\xi_1$ :

$$0 \leq \xi_1 \leq \xi_1^{max}(t, \varphi_0, \psi_0), \quad (46)$$

where  $\xi_1^{max}(t, \varphi_0, \psi_0)$  is a constant uniquely determined for  $\varphi_0$  and  $\psi_0$  by

$$\begin{cases} 0 \leq \xi_1^{max}(t, \varphi_0, \psi_0) < \varphi_0 \\ 1 + w_1(\xi_1^{max}(t, \varphi_0, \psi_0)) = C_1(\xi_1^{max}(t, \varphi_0, \psi_0)). \end{cases} \quad (47)$$

We remark that this can be obtained by a similar geometric method to the previous case: consider the intersection of the line  $l_1$  and the line  $x = -\frac{1}{t}$  in turn (see FIG. 2). Overall, we have demonstrated that  $\xi_1$  for  $\tilde{A}_1^t$  satisfies

$$\xi_1^{min}(t, \varphi_0, \psi_0) \leq \xi_1 \leq \xi_1^{max}(t, \varphi_0, \psi_0), \quad (48)$$

where  $\xi_1^{min}(t, \varphi_0, \psi_0)$  and  $\xi_1^{max}(t, \varphi_0, \psi_0)$  are obtained thorough (45) and (47) respectively. Note that  $\xi_1^{min}(t, \varphi_0, \psi_0)$  and  $\xi_1^{max}(t, \varphi_0, \psi_0)$  depend continuously on  $t$  (and  $\varphi_1, \varphi_2$  through  $\varphi_0$  and  $\psi_0$ ).

Similarly, we consider a binary observable  $\tilde{A}_2^t$  on  $\mathcal{S}_D$  which coincides with  $A_D^t$  in  $\mathcal{S}_1$ , and focus on its effect  $\tilde{A}_2^t(+)$ . We define parameters  $\mathbf{v}_2 \in D$  and  $\xi_2 \in [-\pi, \pi)$  as

$$\begin{aligned} \mathbf{v}_2 &= \sin \xi_2 \mathbf{x} + \cos \xi_2 \mathbf{y} \\ &= \operatorname{argmax}_{\mathbf{v} \in D} \operatorname{tr}[\tilde{A}_2^t(+) \rho_{\mathbf{v}}]. \end{aligned} \quad (49)$$

$\tilde{A}_2^t(+)$  is represented as

$$\tilde{A}_2^t(+) = \frac{1}{2} ((1 + w_2(\xi_2)) \mathbb{1} + \mathbf{m}_2(\xi_2)) \quad (50)$$

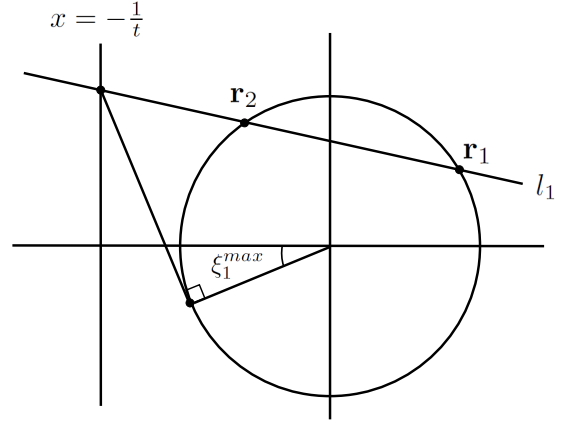


FIG. 2: Geometric description of determining  $\xi_1^{max}$ .

with

$$\mathbf{m}_2(\xi_2) = C_2(\xi_2) \mathbf{v}_2 \quad (0 \leq C_2(\xi_2) \leq 1).$$

(29) becomes

$$\begin{aligned} \frac{1}{2} + \frac{t}{2} \cos\left(\frac{\pi}{2} - \varphi_1\right) \\ = \frac{1 + x_2(\xi_2)}{2} + \frac{C_2(\xi_2)}{2} \cos\left(\frac{\pi}{2} - \varphi_1 - \xi_1\right), \end{aligned} \quad (51)$$

$$\begin{aligned} \frac{1}{2} + \frac{t}{2} \cos\left(\frac{\pi}{2} - \varphi_2\right) \\ = \frac{1 + x_2(\xi_2)}{2} + \frac{C_2(\xi_2)}{2} \cos\left(\frac{\pi}{2} - \varphi_2 - \xi_1\right), \end{aligned}$$

so defining  $\overline{\varphi}_1 := \frac{\pi}{2} - \varphi_1$  and  $\overline{\varphi}_2 := \frac{\pi}{2} - \varphi_2$ , we can obtain similarly to (30) and (31)

$$C_2(\xi_2) = \frac{t \sin \overline{\varphi}_0}{\sin(\overline{\varphi}_0 - \xi_1)}, \quad (52)$$

$$w_2(\xi_2) = \frac{-t \cos \psi_0 \sin \xi_2}{\sin(\overline{\varphi}_0 - \xi_2)}, \quad (53)$$

where  $\overline{\varphi}_0 := \frac{\overline{\varphi}_1 + \overline{\varphi}_2}{2} = \frac{\pi}{2} - \varphi_0$ . It follows that properties of  $\tilde{A}_2^t$  can be obtained just by replacing  $\xi_1$  and  $\varphi_0$  exhibited in the argument for  $\tilde{A}_1^t$  by  $\xi_2$  and  $\overline{\varphi}_0$  respectively. Remark that  $0 < \overline{\varphi}_0 < \frac{\pi}{2}$  holds similarly to  $\varphi_0$ , and that the change  $\psi_0 \rightarrow \overline{\psi}_0 := \frac{\overline{\varphi}_2 - \overline{\varphi}_1}{2} = -\psi_0$  does not affect the equations above, so we dismiss it. From (52) and (53), we have

$$\xi_2^{min}(t, \varphi_0, \psi_0) \leq \xi_2 \leq \xi_2^{max}(t, \varphi_0, \psi_0), \quad (54)$$

where

$$\begin{aligned} \xi_2^{min}(t, \varphi_0, \psi_0) &= \xi_1^{min}(t, \overline{\varphi}_0, \psi_0) \\ &= \xi_1^{min}\left(t, \frac{\pi}{2} - \varphi_0, \psi_0\right), \end{aligned} \quad (55)$$

and

$$\begin{aligned} \xi_2^{max}(t, \varphi_0, \psi_0) &= \xi_1^{max}(t, \overline{\varphi}_0, \psi_0) \\ &= \xi_1^{max}\left(t, \frac{\pi}{2} - \varphi_0, \psi_0\right), \end{aligned} \quad (56)$$

which satisfy

$$\begin{cases} -\frac{\pi}{2} + \varphi_0 < \xi_2^{\min}(t, \varphi_0, \psi_0) \leq 0 \\ 1 - w_2(\xi_2^{\min}(t, \varphi_0, \psi_0)) = C_2(\xi_2^{\min}(t, \varphi_0, \psi_0)) \end{cases} \quad (57)$$

and

$$\begin{cases} 0 \leq \xi_2^{\max}(t, \varphi_0, \psi_0) < \frac{\pi}{2} - \varphi_0 \\ 1 + w_2(\xi_2^{\max}(t, \varphi_0, \psi_0)) = C_2(\xi_2^{\max}(t, \varphi_0, \psi_0)) \end{cases} \quad (58)$$

respectively.

(b) In this part, we shall consider the (in)compatibility of the observables  $\tilde{\mathbf{A}}_1^t$  and  $\tilde{\mathbf{A}}_2^t$  defined in (a) for  $t$  close to  $\frac{1}{\sqrt{2}}$  ( $t \sim \frac{1}{\sqrt{2}}$ ). It is related directly with the  $S_1$ -(in)compatibility of  $\mathbf{A}_D^{tx}$  and  $\mathbf{A}_D^{ty}$  as we have shown in the beginning of this section.

Let us examine the behavior of  $\xi_1^{\min}(t, \varphi_0, \psi_0)$  for  $t \sim \frac{1}{\sqrt{2}}$ . We denote  $\xi_1^{\min}(t = \frac{1}{\sqrt{2}}, \varphi_0, \psi_0)$  and  $h_1(t = \frac{1}{\sqrt{2}}, \varphi_0, \psi_0)$  simply by  $\hat{\xi}_1^{\min}(\varphi_0, \psi_0)$  and  $\hat{h}_1(\varphi_0, \psi_0)$  respectively. The following lemma is useful.

**Lemma 4.** *With  $\varphi_0$  fixed,  $\hat{\xi}_1^{\min}$  is a strictly decreasing function of  $\psi_0$ .*

*Proof.* The claim can be observed to hold by a geometric consideration in terms of FIG. 1. In fact, increasing  $\psi_0$  with  $\varphi_0$  fixed corresponds to moving the line  $l_1$  down with its inclination fixed. The movement makes  $h_1$  (or  $\hat{h}_1$ ) and hence  $\xi_1^{\min}$  (or  $\hat{\xi}_1^{\min}$ ) smaller, which proves the claim. Here we show an analytic proof of this fact. We can see from (38) and (42) that

$$\sqrt{2} \cos \hat{\xi}_1^{\min} + \hat{h}_1 \sin \hat{\xi}_1^{\min} = 1, \quad (59)$$

i.e.

$$\hat{h}_1 = \frac{1}{\sin \hat{\xi}_1^{\min}} \left( 1 - \sqrt{2} \cos \hat{\xi}_1^{\min} \right)$$

holds (note that  $\sin \hat{\xi}_1^{\min} \neq 0$  because  $\sin \hat{\xi}_1^{\min} = 0$  contradicts (59)). Because

$$\frac{d\hat{h}_1}{d\hat{\xi}_1^{\min}} = \frac{1}{(\sin \hat{\xi}_1^{\min})^2} \left( \sqrt{2} - \cos \hat{\xi}_1^{\min} \right) > 0,$$

and  $\hat{h}_1 = \frac{1}{\sin \varphi_0} (\cos \psi_0 - \sqrt{2} \cos \varphi_0)$  is a decreasing function of  $\psi_0$ , the claim follows.  $\square$

From this lemma, it follows that

$$\hat{\xi}_1^{\min}(\varphi_0, \psi_0) < \lim_{\psi_0 \rightarrow +0} \hat{\xi}_1^{\min}(\varphi_0, \psi_0) =: \Xi_1^{\min}(\varphi_0), \quad (60)$$

and

$$\hat{\xi}_2^{\min}(\varphi_0, \psi_0) < \Xi_2^{\min}(\varphi_0) \quad (61)$$

hold for all  $\varphi_0 \in (0, \frac{\pi}{2})$  and  $\psi_0 \in (0, \frac{\pi}{2})$ , where

$$\begin{aligned} \hat{\xi}_2^{\min}(\varphi_0, \psi_0) &:= \xi_2^{\min} \left( t = \frac{1}{\sqrt{2}}, \varphi_0, \psi_0 \right) \\ &= \hat{\xi}_1^{\min} \left( \frac{\pi}{2} - \varphi_0, \psi_0 \right), \\ \Xi_2^{\min}(\varphi_0) &:= \lim_{\psi_0 \rightarrow +0} \hat{\xi}_2^{\min}(\varphi_0, \psi_0) \\ &= \Xi_1^{\min} \left( \frac{\pi}{2} - \varphi_0 \right). \end{aligned} \quad (62)$$

We can prove the following lemma.

**Lemma 5.**

$$\Xi_1^{\min}(\varphi_0) + \Xi_2^{\min}(\varphi_0) \leq -\frac{\pi}{2}$$

holds for all  $0 < \varphi_0 < \frac{\pi}{2}$ .

*Proof.* Let us define

$$\begin{aligned} H_1(\varphi_0) &:= \lim_{\psi_0 \rightarrow +0} \hat{h}_1(\varphi_0, \psi_0) \\ &= \lim_{\psi_0 \rightarrow +0} h_1 \left( t = \frac{1}{\sqrt{2}}, \varphi_0, \psi_0 \right) \\ &= \frac{1}{\sin \varphi_0} \left( 1 - \sqrt{2} \cos \varphi_0 \right). \end{aligned}$$

It holds similarly to (59) that

$$\sqrt{2} \cos \Xi_1^{\min} + H_1 \sin \Xi_1^{\min} = 1. \quad (63)$$

Hence, together with  $\sin^2 \Xi_1^{\min} + \cos^2 \Xi_1^{\min} = 1$ , we can obtain

$$\cos \Xi_1^{\min} = \frac{1}{\sqrt{2}} \cdot \frac{2 + H_1 \sqrt{2H_1^2 + 2}}{H_1^2 + 2}, \quad (64)$$

or its more explicit form

$$\cos \Xi_1^{\min} = \frac{1}{\sqrt{2}} \cdot \frac{4 - 3\sqrt{2} \cos \varphi_0}{3 - 2\sqrt{2} \cos \varphi_0}. \quad (65)$$

It results in

$$\Xi_1^{\min}(\varphi_0) = -\arccos \left( \frac{1}{\sqrt{2}} \cdot \frac{4 - 3\sqrt{2} \cos \varphi_0}{3 - 2\sqrt{2} \cos \varphi_0} \right), \quad (66)$$

where we follow the convention that  $\arccos: [-1, 1] \rightarrow [0, \pi]$ , and thus  $\Xi_1^{\min} \in (-\pi + \varphi_0, 0]$  is obtained through  $-\arccos: [-1, 1] \rightarrow [-\pi, 0]$ . Because

$$\frac{d}{d\varphi_0} \left( \frac{1}{\sqrt{2}} \cdot \frac{4 - 3\sqrt{2} \cos \varphi_0}{3 - 2\sqrt{2} \cos \varphi_0} \right) = \frac{\sin \varphi_0}{(3 - 2\sqrt{2} \cos \varphi_0)^2},$$

and

$$\begin{aligned} \sqrt{1 - \left( \frac{1}{\sqrt{2}} \cdot \frac{4 - 3\sqrt{2} \cos \varphi_0}{3 - 2\sqrt{2} \cos \varphi_0} \right)^2} &= \sqrt{\left( \frac{\sin \varphi_0}{3 - 2\sqrt{2} \cos \varphi_0} \right)^2} \\ &= \frac{\sin \varphi_0}{3 - 2\sqrt{2} \cos \varphi_0}, \end{aligned}$$

we can observe that

$$\begin{aligned} \frac{d\Xi_1^{\min}}{d\varphi_0} &= \left( \frac{\sin \varphi_0}{3 - 2\sqrt{2} \cos \varphi_0} \right)^{-1} \cdot \frac{\sin \varphi_0}{(3 - 2\sqrt{2} \cos \varphi_0)^2} \\ &= \frac{1}{3 - 2\sqrt{2} \cos \varphi_0}, \end{aligned}$$

and

$$\frac{d^2\Xi_1^{\min}}{d\varphi_0^2} = \frac{-2\sqrt{2} \sin \varphi_0}{(3 - 2\sqrt{2} \cos \varphi_0)^2} < 0, \quad (67)$$

which means  $\Xi_1^{\min}$  is concave. Therefore, for any  $\varphi_0 \in (0, \frac{\pi}{2})$ , the concavity results in

$$\begin{aligned} \frac{1}{2}\Xi_1^{\min}(\varphi_0) + \frac{1}{2}\Xi_2^{\min}(\varphi_0) &= \frac{1}{2}\Xi_1^{\min}(\varphi_0) + \frac{1}{2}\Xi_1^{\min}\left(\frac{\pi}{2} - \varphi_0\right) \\ &\leq \Xi_1^{\min}\left(\frac{1}{2}\varphi_0 + \frac{1}{2}\left(\frac{\pi}{2} - \varphi_0\right)\right) \\ &= \Xi_1^{\min}\left(\frac{\pi}{4}\right). \end{aligned}$$

Since we can see from (66) that  $\Xi_1^{\min}(\frac{\pi}{4}) = -\frac{\pi}{4}$ ,

$$\Xi_1^{\min}(\varphi_0) + \Xi_2^{\min}(\varphi_0) \leq -\frac{\pi}{2}$$

holds for any  $\varphi_0 \in (0, \frac{\pi}{2})$ .  $\square$

According to Lemma 4 and Lemma 5,

$$\begin{aligned} \widehat{\xi}_1^{\min}(\varphi_0, \psi_0) + \widehat{\xi}_2^{\min}(\varphi_0, \psi_0) &< \Xi_1^{\min}(\varphi_0) + \Xi_2^{\min}(\varphi_0) \\ &\leq -\frac{\pi}{2}, \end{aligned}$$

that is,

$$\xi_1^{\min}\left(t = \frac{1}{\sqrt{2}}, \varphi_0, \psi_0\right) + \xi_2^{\min}\left(t = \frac{1}{\sqrt{2}}, \varphi_0, \psi_0\right) < -\frac{\pi}{2}$$

holds for any  $\varphi_0$  and  $\psi_0$  (i.e. for any  $\varphi_1$  and  $\varphi_2$ ). However, we cannot conclude that

$$\xi_1^{\min}(t, \varphi_0, \psi_0) + \xi_2^{\min}(t, \varphi_0, \psi_0) \leq -\frac{\pi}{2} \quad (68)$$

holds for  $t \sim \frac{1}{\sqrt{2}}$ : it may fail when

$$\begin{aligned} \sup_{\varphi_0, \psi_0} \left[ \xi_1^{\min}\left(t = \frac{1}{\sqrt{2}}, \varphi_0, \psi_0\right) \right. \\ \left. + \xi_2^{\min}\left(t = \frac{1}{\sqrt{2}}, \varphi_0, \psi_0\right) \right] = -\frac{\pi}{2}. \end{aligned}$$

On the other hand, because we can observe similarly to Lemma 4 that  $\xi_1^{\min}$  is a strictly decreasing function of  $\psi_0$ , it is anticipated that (68) holds for  $t \sim \frac{1}{\sqrt{2}}$  and for  $\psi_0$  sufficiently close to  $\frac{\pi}{2}$ . In fact, for  $\psi_0 \in [\frac{\pi}{4}, \frac{\pi}{2})$ , we can prove the following proposition.

**Proposition 5.** *There exists a constant  $C < -\frac{\pi}{2}$  such that*

$$\widehat{\xi}_1^{\min}(\varphi_0, \psi_0) + \widehat{\xi}_2^{\min}(\varphi_0, \psi_0) < C,$$

i.e.

$$\xi_1^{\min}\left(t = \frac{1}{\sqrt{2}}, \varphi_0, \psi_0\right) + \xi_2^{\min}\left(t = \frac{1}{\sqrt{2}}, \varphi_0, \psi_0\right) < C,$$

holds for all  $\psi_0 \in [\frac{\pi}{4}, \frac{\pi}{2})$  and  $\varphi_0 \in (0, \frac{\pi}{2})$ .

*Proof.* Because

$$\begin{aligned} \widehat{\xi}_1^{\min}(\varphi_0, \psi_0) + \widehat{\xi}_2^{\min}(\varphi_0, \psi_0) &= \widehat{\xi}_1^{\min}(\varphi_0, \psi_0) + \widehat{\xi}_1^{\min}\left(\frac{\pi}{2} - \varphi_0, \psi_0\right), \end{aligned}$$

we can assume without loss of generality that  $0 < \varphi_0 \leq \frac{\pi}{4}$ . Due to Lemma 4, it holds for any  $\psi_0 \in [\frac{\pi}{4}, \frac{\pi}{2})$  that

$$\begin{aligned} \widehat{\xi}_1^{\min}(\varphi_0, \psi_0) &\leq \widehat{\xi}_1^{\min}\left(\varphi_0, \psi_0 = \frac{\pi}{4}\right), \\ \widehat{\xi}_1^{\min}\left(\frac{\pi}{2} - \varphi_0, \psi_0\right) &\leq \widehat{\xi}_1^{\min}\left(\frac{\pi}{2} - \varphi_0, \psi_0 = \frac{\pi}{4}\right). \end{aligned} \quad (69)$$

Let us denote  $\widehat{\xi}_1^{\min}(\varphi_0, \psi_0 = \frac{\pi}{4})$  simply by  $\widetilde{\Xi}_1^{\min}(\varphi_0)$ . In order to investigate  $\widetilde{\Xi}_1^{\min}(\varphi_0)$  and  $\widetilde{\Xi}_1^{\min}(\frac{\pi}{2} - \varphi_0)$ , we have to recall (59). Similarly to (63) and (64) in the proof of Lemma 5, it results in

$$\cos \widetilde{\Xi}_1^{\min} = \frac{1}{\sqrt{2}} \cdot \frac{2 + \widetilde{H}_1 \sqrt{2\widetilde{H}_1^2 + 2}}{\widetilde{H}_1^2 + 2}, \quad (70)$$

where

$$\begin{aligned} \widetilde{H}_1(\varphi_0) &= \widehat{h}_1\left(\varphi_0, \psi_0 = \frac{\pi}{4}\right) \\ &= \frac{1}{\sin \varphi_0} \left( \frac{1}{\sqrt{2}} - \sqrt{2} \cos \varphi_0 \right). \end{aligned} \quad (71)$$

Note that in this case we cannot apply a similar method to the one in Lemma 5 because  $\widetilde{\Xi}_1^{\min}$  does not have a clear form like (66). Alternatively, we focus on the following monotone relations between  $\widetilde{\Xi}_1^{\min}$ ,  $\widetilde{H}_1$ , and  $\varphi_0$  (referring to the proof of Lemma 4 may be helpful):

$$\frac{d\widetilde{\Xi}_1^{\min}}{d\widetilde{H}_1} > 0, \quad \frac{d\widetilde{H}_1}{d\varphi_0} > 0 \quad \left( \text{thus } \frac{d\widetilde{\Xi}_1^{\min}}{d\varphi_0} > 0 \right). \quad (72)$$

From these relations, it can be seen that our restriction  $0 < \varphi_0 \leq \frac{\pi}{4}$  is equivalent to the condition  $\widetilde{H}_1 \leq 1 - \sqrt{2}$  since  $\widetilde{H}_1(0) = -\infty$  and  $\widetilde{H}_1(\frac{\pi}{4}) = 1 - \sqrt{2}$ . The claim of the proposition can be shown easily when  $\widetilde{H}_1 \leq -1$  (or  $0 < \varphi_0 \leq \varphi^* := \arccos \frac{2+\sqrt{10}}{6}$ , where  $\widetilde{H}_1(\varphi^*) = -1$ ). In fact,

$$\widetilde{\Xi}_1^{\min}(\varphi_0) \leq \widetilde{\Xi}_1^{\min}(\varphi^*) = -\frac{\pi}{2}$$

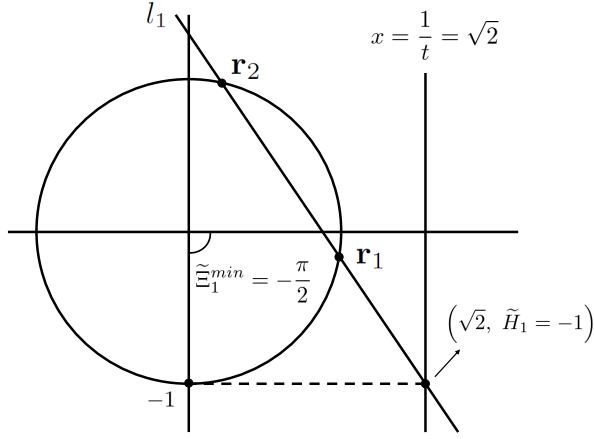


FIG. 3: Geometric description of  $\tilde{\Xi}_1^{min}$ . It can be observed that  $\tilde{\Xi}_1^{min} = -\frac{\pi}{2}$  when  $\tilde{H}_1 = -1$ .

and

$$\tilde{\Xi}_1^{min} \left( \frac{\pi}{2} - \varphi_0 \right) < \tilde{\Xi}_1^{min} \left( \frac{\pi}{2} \right) = -\arccos \frac{2\sqrt{2} + \sqrt{3}}{5}$$

hold (see FIG. 3 and (70)), and thus we can conclude

$$\tilde{\Xi}_1^{min}(\varphi_0) + \tilde{\Xi}_1^{min} \left( \frac{\pi}{2} - \varphi_0 \right) < C_1,$$

where

$$C_1 = -\frac{\pi}{2} - \arccos \frac{2\sqrt{2} + \sqrt{3}}{5} \left( < -\frac{\pi}{2} \right).$$

When  $-1 < \tilde{H}_1 \leq 1 - \sqrt{2}$  (or  $\varphi^* < \varphi_0 \leq \frac{\pi}{4}$ ), we need a bit complicated evaluations. It holds similarly to the previous calculations that

$$\begin{aligned} \tilde{\Xi}_1^{min}(\varphi_0) &\leq \tilde{\Xi}_1^{min} \left( \frac{\pi}{4} \right), \\ \tilde{\Xi}_1^{min} \left( \frac{\pi}{2} - \varphi_0 \right) &< \tilde{\Xi}_1^{min} \left( \frac{\pi}{2} - \varphi^* \right). \end{aligned}$$

Since

$$\tilde{\Xi}_1^{min} = -\frac{\pi}{4} \iff \tilde{H}_1 = 0 \iff \varphi_0 = \frac{\pi}{3},$$

$\tilde{\Xi}_1^{min}(\frac{\pi}{4}) < -\frac{\pi}{4} = \tilde{\Xi}_1^{min}(\frac{\pi}{3})$  holds due to the monotone relations (72). On the other hand, we have

$$\begin{aligned} \cos \varphi^* - \cos \frac{\pi}{6} &= \frac{2 + \sqrt{10}}{6} - \frac{\sqrt{3}}{2} \\ &= -0.0056... < 0, \end{aligned}$$

that is,

$$\varphi^* > \frac{\pi}{6}.$$

It follows that  $\frac{\pi}{2} - \varphi^* < \frac{\pi}{3}$ , and thus  $\tilde{\Xi}_1^{min}(\frac{\pi}{2} - \varphi_0) < -\frac{\pi}{4}$ . Therefore, we can conclude also in this case

$$\tilde{\Xi}_1^{min}(\varphi_0) + \tilde{\Xi}_1^{min} \left( \frac{\pi}{2} - \varphi_0 \right) < C_2,$$

where

$$C_2 = \tilde{\Xi}_1^{min} \left( \frac{\pi}{4} \right) + \tilde{\Xi}_1^{min} \left( \frac{\pi}{2} - \varphi^* \right) \left( < -\frac{\pi}{2} \right).$$

Overall, we have obtained

$$\tilde{\Xi}_1^{min}(\varphi_0, \psi_0) + \tilde{\Xi}_2^{min}(\varphi_0, \psi_0) < \max\{C_1, C_2\} \left( < -\frac{\pi}{2} \right)$$

for all  $\varphi_0 \in (0, \frac{\pi}{2})$  and  $\psi_0 \in [\frac{\pi}{4}, \frac{\pi}{2})$ .  $\square$

By virtue of this proposition, for  $t$  sufficiently close to  $\frac{1}{\sqrt{2}}$ ,

$$\xi_1^{min}(t, \varphi_0, \psi_0) + \xi_2^{min}(t, \varphi_0, \psi_0) \leq -\frac{\pi}{2}$$

follows from the continuity of  $\xi_1^{min}$  and  $\xi_2^{min}$  with respect to  $t$  when  $\frac{\pi}{4} \leq \psi_0 < \frac{\pi}{2}$ . It means that there always exist  $\xi_1^* \geq \xi_1^{min}$  and  $\xi_2^* \geq \xi_2^{min}$  for such  $t$  and for any  $\varphi_1$  and  $\varphi_2$  satisfying  $\xi_1^* + \xi_2^* = -\frac{\pi}{2}$ . For these  $\xi_1^*$  and  $\xi_2^*$ ,  $\mathbf{v}_1 = -\mathbf{v}_2$  holds, and thus  $\tilde{\mathbf{A}}_1^t$  and  $\tilde{\mathbf{A}}_2^t$  are compatible, i.e.  $\mathbf{A}_D^{tx}$  and  $\mathbf{A}_D^{ty}$  are  $\mathcal{S}_1$ -compatible.

On the other hand, when  $0 < \psi_0 < \frac{\pi}{4}$ , it may not hold for  $t \sim \frac{1}{\sqrt{2}}$  that  $\xi_1^{min}(t, \varphi_0, \psi_0) + \xi_2^{min}(t, \varphi_0, \psi_0) \leq -\frac{\pi}{2}$ , and thus we cannot apply the same argument. Nevertheless, we can demonstrate that there exist  $\xi_1$  and  $\xi_2$  such that  $\tilde{\mathbf{A}}_1^t$  and  $\tilde{\mathbf{A}}_2^t$  are compatible even when  $0 < \psi_0 < \frac{\pi}{4}$ . To see this, let us assume  $0 < \psi_0 < \frac{\pi}{4}$  and apply the necessary and sufficient condition for (in)compatibility. According to the result proved in [30–32],  $\tilde{\mathbf{A}}_1^t$  and  $\tilde{\mathbf{A}}_2^t$  with (27) and (50) respectively are compatible if and only if

$$\begin{aligned} (1 - F_1^2 - F_2^2) \left( 1 - \frac{w_1^2}{F_1^2} - \frac{w_2^2}{F_2^2} \right) \\ \leq (\mathbf{m}_1 \cdot \mathbf{m}_2 - w_1 w_2)^2 \end{aligned} \quad (73)$$

holds, where

$$F_1 := \frac{1}{2} \left( \sqrt{(1 + w_1)^2 - C_1^2} + \sqrt{(1 - w_1)^2 - C_1^2} \right), \quad (74)$$

$$F_2 := \frac{1}{2} \left( \sqrt{(1 + w_2)^2 - C_2^2} + \sqrt{(1 - w_2)^2 - C_2^2} \right). \quad (75)$$

For  $\xi_1^{min}$  and  $\xi_2^{min}$ , since it holds that

$$1 - w_1(\xi_1^{min}) = C_1(\xi_1^{min}), \quad (76)$$

$$1 - w_2(\xi_2^{min}) = C_2(\xi_2^{min}), \quad (77)$$

they become

$$F_1 = \sqrt{w_1(\xi_1^{min})}, \quad F_2 = \sqrt{w_2(\xi_2^{min})}. \quad (78)$$

Therefore, (73) can be rewritten as

$$\begin{aligned} & [(1 - \sin(\xi_1^{\min} + \xi_2^{\min}))w_1(\xi_1^{\min})w_2(\xi_2^{\min}) - (1 + \sin(\xi_1^{\min} + \xi_2^{\min}))(1 - w_1(\xi_1^{\min}) - w_2(\xi_2^{\min}))] \\ & [(1 - w_1(\xi_1^{\min}))(1 - w_2(\xi_2^{\min}))(1 - \sin(\xi_1^{\min} + \xi_2^{\min}))] \geq 0. \end{aligned} \quad (79)$$

If  $1 - \sin(\xi_1^{\min} + \xi_2^{\min}) = 0$ , then (79) holds, that is,  $\tilde{A}_1^t$  and  $\tilde{A}_2^t$  for  $\xi_1^{\min}$  and  $\xi_2^{\min}$  respectively are compatible. Therefore, we hereafter assume  $1 - \sin(\xi_1^{\min} + \xi_2^{\min}) > 0$ , and rewrite (79) as (note that  $0 < w_1(\xi_1^{\min}) < 1$ ,  $0 < w_2(\xi_2^{\min}) < 1$ )

$$\begin{aligned} & (1 + \sin(\xi_1^{\min} + \xi_2^{\min}))(1 - w_1(\xi_1^{\min}) - w_2(\xi_2^{\min})) \\ & \leq (1 - \sin(\xi_1^{\min} + \xi_2^{\min}))w_1(\xi_1^{\min})w_2(\xi_2^{\min}). \end{aligned} \quad (80)$$

In other words,  $\tilde{A}_1^t$  and  $\tilde{A}_2^t$  with respect to  $\xi_1^{\min}$  and  $\xi_2^{\min}$  are incompatible if and only if

$$\begin{aligned} & (1 + \sin(\xi_1^{\min} + \xi_2^{\min}))(1 - w_1(\xi_1^{\min}) - w_2(\xi_2^{\min})) \\ & > (1 - \sin(\xi_1^{\min} + \xi_2^{\min}))w_1(\xi_1^{\min})w_2(\xi_2^{\min}) \end{aligned} \quad (81)$$

holds. In order to investigate whether (81) holds, it is helpful to introduce a function  $Z$  defined as

$$\begin{aligned} Z(t, \varphi_0, \psi_0) & := [1 + \sin(\xi_1^{\min}(t, \varphi_0, \psi_0) + \xi_2^{\min}(t, \varphi_0, \psi_0))] [1 + w_1(\xi_1^{\min}(t, \varphi_0, \psi_0)) + w_2(\xi_2^{\min}(t, \varphi_0, \psi_0))] \\ & - [1 - \sin(\xi_1^{\min}(t, \varphi_0, \psi_0) + \xi_2^{\min}(t, \varphi_0, \psi_0))] w_1(\xi_1^{\min}(t, \varphi_0, \psi_0))w_2(\xi_2^{\min}(t, \varphi_0, \psi_0)). \end{aligned} \quad (82)$$

Because

$$\begin{aligned} & (1 + \sin(\xi_1^{\min} + \xi_2^{\min}))(1 - w_1(\xi_1^{\min}) - w_2(\xi_2^{\min})) \\ & < (1 + \sin(\xi_1^{\min} + \xi_2^{\min}))(1 + w_1(\xi_1^{\min}) + w_2(\xi_2^{\min})), \\ & Z(t, \varphi_0, \psi_0) > 0 \end{aligned} \quad (83)$$

holds if  $\tilde{A}_1^t$  and  $\tilde{A}_2^t$  with respect to  $\xi_1^{\min}$  and  $\xi_2^{\min}$  are incompatible. Let us focus on the case when  $t = \frac{1}{\sqrt{2}}$  (i.e.  $\xi_1^{\min} = \hat{\xi}_1^{\min}$ ). If a pair  $(\varphi_0, \psi_0)$  satisfies  $\hat{\xi}_1^{\min}(\varphi_0, \psi_0) \leq -\frac{\pi}{2}$  or  $\hat{\xi}_2^{\min}(\varphi_0, \psi_0) \leq -\frac{\pi}{2}$ , then

$$\hat{\xi}_1^{\min}(\varphi_0, \psi_0) + \hat{\xi}_2^{\min}(\varphi_0, \psi_0) < C$$

with

$$\begin{aligned} C & = -\frac{\pi}{2} + \lim_{\substack{\varphi_0 \rightarrow \frac{\pi}{2}-0 \\ \psi_0 \rightarrow +0}} \hat{\xi}_1^{\min}(\varphi_0, \psi_0) \\ & = -\frac{\pi}{2} - \arccos\left(\frac{2\sqrt{2}}{3}\right) \\ & < -\frac{\pi}{2} \end{aligned}$$

holds due to similar monotone relations to (72) between  $\varphi_0, \psi_0$ , and  $\hat{\xi}_1^{\min}$  (remember that  $\hat{\xi}_2^{\min}(\varphi_0, \psi_0) = \hat{\xi}_1^{\min}(\frac{\pi}{2} - \varphi_0, \psi_0)$ ). Therefore, in this case, we can apply the same argument as Proposition 5, which results in the compatibility of  $\tilde{A}_1^t$  and  $\tilde{A}_2^t$  for  $t \sim \frac{1}{\sqrt{2}}$ . On the other hand, let us examine the case when  $(\varphi_0, \psi_0)$  satisfies  $\psi_0 \in (0, \frac{\pi}{4})$ , and  $\hat{\xi}_1^{\min}(\varphi_0, \psi_0) > -\frac{\pi}{2}$  and  $\hat{\xi}_2^{\min}(\varphi_0, \psi_0) > -\frac{\pi}{2}$ .

Because  $\psi_0 \in (0, \frac{\pi}{4})$ , we obtain for general  $t$  (see (31))

$$\begin{aligned} w_1(\xi_1^{\min}) & > \frac{t}{\sqrt{2}} \frac{\sin \xi_1^{\min}}{\sin(\varphi_0 - \xi_1^{\min})} \\ & \geq \frac{t}{\sqrt{2}} (-\sin \xi_1^{\min}). \end{aligned} \quad (84)$$

For  $t = \frac{1}{\sqrt{2}}$ , since

$$-\frac{\pi}{2} < \hat{\xi}_1^{\min}(\varphi_0, \psi_0) < \lim_{\substack{\varphi_0 \rightarrow \frac{\pi}{2}-0 \\ \psi_0 \rightarrow +0}} \hat{\xi}_1^{\min}(\varphi_0, \psi_0),$$

it gives a bound

$$w_1(\xi_1^{\min}) > \frac{1}{2} \sin \hat{\xi}_0, \quad (85)$$

where we define

$$\begin{aligned} \hat{\xi}_0 & = - \lim_{\substack{\varphi_0 \rightarrow \frac{\pi}{2}-0 \\ \psi_0 \rightarrow +0}} \hat{\xi}_1^{\min}(\varphi_0, \psi_0) \\ & = \arccos\left(\frac{2\sqrt{2}}{3}\right). \end{aligned}$$

Let  $\varepsilon$  be a positive constant satisfying  $\varepsilon < \frac{1}{16}(\sin \hat{\xi}_0)^2$ . Due to the continuity of sine, there exists a positive constant  $\delta$  such that  $\sin x \in (-1, -1 + \varepsilon)$  whenever  $x \in (-\frac{\pi}{2} - \delta, -\frac{\pi}{2})$ . If  $(\varphi_0, \psi_0)$  satisfies  $\hat{\xi}_1^{\min}(\varphi_0, \psi_0) + \hat{\xi}_2^{\min}(\varphi_0, \psi_0) \leq -\frac{\pi}{2} - \delta$ , then it again leads to the same argument as Proposition 5, and we can see that  $\tilde{A}_1^t$  and  $\tilde{A}_2^t$  for this  $(\varphi_0, \psi_0)$  are compatible. Conversely, if  $(\varphi_0, \psi_0)$

satisfies  $-\frac{\pi}{2} - \delta < \widehat{\xi}_1^{\min}(\varphi_0, \psi_0) + \widehat{\xi}_2^{\min}(\varphi_0, \psi_0) < -\frac{\pi}{2}$  (remember Lemma 5), then

$$-1 < \sin(\widehat{\xi}_1^{\min}(\varphi_0, \psi_0) + \widehat{\xi}_2^{\min}(\varphi_0, \psi_0)) < -1 + \varepsilon$$

follows from the definition of  $\delta$ . Therefore, by virtue of (82), we have

$$\begin{aligned} Z\left(t = \frac{1}{\sqrt{2}}, \varphi_0, \psi_0\right) &= \left[1 + \sin(\widehat{\xi}_1^{\min}(\varphi_0, \psi_0) + \widehat{\xi}_2^{\min}(\varphi_0, \psi_0))\right] \left[1 + w_1(\widehat{\xi}_1^{\min}(\varphi_0, \psi_0)) + w_2(\widehat{\xi}_2^{\min}(\varphi_0, \psi_0))\right] \\ &\quad - \left[1 - \sin(\widehat{\xi}_1^{\min}(\varphi_0, \psi_0) + \widehat{\xi}_2^{\min}(\varphi_0, \psi_0))\right] w_1(\widehat{\xi}_1^{\min}(\varphi_0, \psi_0))w_2(\widehat{\xi}_2^{\min}(\varphi_0, \psi_0)) \\ &< \varepsilon \left[1 + w_1(\widehat{\xi}_1^{\min}(\varphi_0, \psi_0)) + w_2(\widehat{\xi}_2^{\min}(\varphi_0, \psi_0))\right] - (2 - \varepsilon)w_1(\widehat{\xi}_1^{\min}(\varphi_0, \psi_0))w_2(\widehat{\xi}_2^{\min}(\varphi_0, \psi_0)) \\ &= \varepsilon \left[1 + w_1(\widehat{\xi}_1^{\min}(\varphi_0, \psi_0))\right] \left[1 + w_2(\widehat{\xi}_2^{\min}(\varphi_0, \psi_0))\right] - 2w_1(\widehat{\xi}_1^{\min}(\varphi_0, \psi_0))w_2(\widehat{\xi}_2^{\min}(\varphi_0, \psi_0)) \\ &< 4\varepsilon - 2w_1(\widehat{\xi}_1^{\min}(\varphi_0, \psi_0))w_2(\widehat{\xi}_2^{\min}(\varphi_0, \psi_0)). \end{aligned}$$

Because

$$\begin{aligned} 4\varepsilon - 2w_1(\widehat{\xi}_1^{\min}(\varphi_0, \psi_0))w_2(\widehat{\xi}_2^{\min}(\varphi_0, \psi_0)) \\ < \frac{1}{4}(\sin \widehat{\xi}_0)^2 - \frac{1}{2}(\sin \widehat{\xi}_0)^2 = -\frac{1}{4}(\sin \widehat{\xi}_0)^2, \end{aligned}$$

it holds that

$$Z\left(t = \frac{1}{\sqrt{2}}, \varphi_0, \psi_0\right) < -\frac{1}{4}(\sin \widehat{\xi}_0)^2 < 0.$$

Therefore, for  $t \sim \frac{1}{\sqrt{2}}$ ,  $Z(t, \varphi_0, \psi_0) \leq 0$  holds, that is,  $\widetilde{\mathbf{A}}_1^t$  and  $\widetilde{\mathbf{A}}_2^t$  with respect to  $\xi_1^{\min}$  and  $\xi_2^{\min}$  are compatible. Overall, we have demonstrated that when  $t \sim \frac{1}{\sqrt{2}}$ , there exist compatible observables  $\widetilde{\mathbf{A}}_1^t$  and  $\widetilde{\mathbf{A}}_2^t$  for any line  $\mathcal{S}_1 \subset \mathcal{S}_D$  such that they agree with  $\mathbf{A}_D^{tx}$  and  $\mathbf{A}_D^{ty}$  on  $\mathcal{S}_1$  respectively. That is, when  $t \sim \frac{1}{\sqrt{2}}$ ,  $\mathbf{A}_D^{tx}$  and  $\mathbf{A}_D^{ty}$  are  $\mathcal{S}_1$ -compatible for any line  $\mathcal{S}_1 \subset \mathcal{S}_D$ . Therefore, we can conclude that  $\chi_{incomp}(\mathbf{A}_D^{tx}, \mathbf{A}_D^{ty}) = 3$  for  $t \sim \frac{1}{\sqrt{2}}$ , and thus the set  $M$  in (18) is nonempty.

#### Proof of Proposition 4 : Part 2

In this part, we shall show that

$$t'_0 := \inf L = \sup M \in M,$$

where  $L$  and  $M$  are defined in (18). In order to prove this, we will see that if  $t \in L$ , then  $t - \delta \in L$  for sufficiently small  $\delta > 0$ , that is,  $t'_0 \notin L$ .

Let us focus again on a system described by a two-dimensional disk state space  $\mathcal{S}_D$ . It is useful to identify this system with the system of a quantum bit with real coefficients by replacing  $\{\sigma_1, \sigma_2\}$  with  $\{\sigma_3, \sigma_1\}$ . Then, defining  $\mathcal{E}_D$  as the set of all effects on  $\mathcal{S}_D$ , we can see that any  $E \in \mathcal{E}_D$  can be expressed as a real-coefficient positive matrix smaller than  $\mathbb{1}$ . We also define  $O_D(2) \subset \mathcal{E}_D \times \mathcal{E}_D$

as the set of all binary observables on  $\mathcal{S}_D$ , which is isomorphic naturally to  $\mathcal{E}_D$  since a binary observable  $\mathbf{A}$  is completely specified by its effect  $\mathbf{A}(+) \in \mathcal{E}_D$ . With introducing a topology (e.g. norm topology) on  $\mathcal{E}_D$ , it also can be observed that  $O_D(2)$  is homeomorphic to  $\mathcal{E}_D$ . Note that because the system is described by finite dimensional matrices, any (natural) topology (norm topology, weak topology, etc.) coincides with each other. For a pair of states  $\{\rho^{r_1}, \rho^{r_2}\}$  in  $\mathcal{S}_D$ , and a binary observable  $\mathbf{A} \in O_D(2)$ , we define a set of observables  $C(\mathbf{A} : \rho^{r_1}, \rho^{r_2})$  as the set of all binary observables  $\widetilde{\mathbf{A}} \in O_D(2)$  such that

$$\begin{aligned} \text{tr}[\rho^{r_1} \widetilde{\mathbf{A}}(\pm)] &= \text{tr}[\rho^{r_1} \mathbf{A}(\pm)], \\ \text{tr}[\rho^{r_2} \widetilde{\mathbf{A}}(\pm)] &= \text{tr}[\rho^{r_2} \mathbf{A}(\pm)]. \end{aligned}$$

It can be confirmed easily that  $C(\mathbf{A} : \rho^{r_1}, \rho^{r_2})$  is closed in  $O_D(2) \simeq \mathcal{E}_D$ . Let us denote by  $O_D(4)$  the set of all observables with four outcomes, which is a compact (i.e. bounded and closed) subset of  $\mathcal{E}_D^4$ . For each  $\mathbf{M} = \{\mathbf{M}(x, y)\} \in O_D(4)$ , we can introduce a pair of binary observables by

$$\pi_1(\mathbf{M}) = \left\{ \sum_y \mathbf{M}(x, y) \right\}_x, \quad \pi_2(\mathbf{M}) = \left\{ \sum_x \mathbf{M}(x, y) \right\}_y.$$

Since  $\pi_j : O_D(4) \rightarrow O_D(2)$  is continuous, the set of all compatible binary observables denoted by

$$JM(2, 2) := \{(\pi_1(\mathbf{M}), \pi_2(\mathbf{M})) \mid \mathbf{M} \in O_D(4)\}$$

is compact in  $O_D(2) \times O_D(2) \simeq \mathcal{E}_D \times \mathcal{E}_D$  as well. As we have seen in the previous part,  $\chi_{incomp}(\mathbf{A}^{tx}, \mathbf{A}^{ty}) = 2$  (i.e.  $\chi_{incomp}(\mathbf{A}_D^{tx}, \mathbf{A}_D^{ty}) = 2$ ) if and only if there exists a pair of vectors  $\mathbf{r}_1, \mathbf{r}_2 \in \partial D$  such that

$$(C(\mathbf{A}_D^{tx} : \rho^{r_1}, \rho^{r_2}) \times C(\mathbf{A}_D^{ty} : \rho^{r_1}, \rho^{r_2})) \cap JM(2 : 2) = \emptyset.$$

Let us examine concrete representations of the sets. Each effect  $E \in \mathcal{E}_D$  is written as  $E = \frac{1}{2}(e_0\mathbb{1} + \mathbf{e} \cdot \sigma) = \frac{1}{2}(e_0\mathbb{1} + e_1\sigma_1 + e_2\sigma_2)$  with  $(e_0, \mathbf{e}) = (e_0, e_1, e_2) \in \mathbb{R}^3$  satisfying  $0 \leq e_0 \pm |\mathbf{e}| \leq 2$ .

If we consider another effect  $F = \frac{1}{2}(f_0\mathbb{1} + \mathbf{f} \cdot \sigma)$ , the operator norm of  $E - F$  is calculated as

$$\|E - F\| = \frac{1}{2}(|e_0 - f_0| + |\mathbf{e} - \mathbf{f}|). \quad (86)$$

We may employ this norm to define a topology on  $\mathcal{E}_D$  and  $O_D(2) \simeq \mathcal{E}_D$ . On the other hand, each state in  $\mathcal{S}_D$  is parameterized as  $\rho^{\mathbf{r}_1} = \frac{1}{2}(\mathbb{1} + x_1\sigma_1 + y_1\sigma_2)$ , where  $\mathbf{r}_1 = (x_1, y_1)$  satisfies  $|\mathbf{r}_1| \leq 1$ . For an effect  $E$  and a state  $\rho^{\mathbf{r}_1}$ , we have  $\text{tr}[\rho^{\mathbf{r}_1}E] = \frac{1}{2}(e_0 + \mathbf{r}_1 \cdot \mathbf{e})$ . In particular, when considering  $A^{t\mathbf{x}}(\pm) = \frac{1}{2}(\mathbb{1} \pm t\sigma_1)$ , a binary observable  $C$  determined by the effect  $C(+) = \frac{1}{2}(c_0\mathbb{1} + \mathbf{c} \cdot \sigma) = \frac{1}{2}(c_0\mathbb{1} + c_1\sigma_1 + c_2\sigma_2)$  satisfies  $C \in C(A^{t\mathbf{x}} : \rho^{\mathbf{r}_1}, \rho^{\mathbf{r}_2})$  if and only if

$$\begin{aligned} \text{tr}[\rho^{\mathbf{r}_1}A^{t\mathbf{x}}(+)] &= \text{tr}[\rho^{\mathbf{r}_1}C(+)], \\ \text{tr}[\rho^{\mathbf{r}_2}A^{t\mathbf{x}}(+)] &= \text{tr}[\rho^{\mathbf{r}_2}C(+)], \end{aligned}$$

i.e.

$$\begin{aligned} 1 + tx_1 &= c_0 + \mathbf{r}_1 \cdot \mathbf{c} = c_0 + x_1c_1 + y_1c_2, \\ 1 + tx_2 &= c_0 + \mathbf{r}_2 \cdot \mathbf{c} = c_0 + x_2c_1 + y_2c_2. \end{aligned}$$

hold, where we set  $\mathbf{r}_2 = (x_2, y_2)$ . The set of their solutions for  $(c_0, \mathbf{c})$  is represented as

$$(c_0, \mathbf{c}) = (1, t, 0) + \lambda' \left( -\frac{x_1y_2 - y_1x_2}{x_1 - x_2}, -\frac{y_1 - y_2}{x_1 - x_2}, 1 \right)$$

with  $\lambda' \in \mathbb{R}$ . Let us define a vector  $\mathbf{n} \in \mathbb{R}^2$  such that

$$(1, \mathbf{n}) \cdot (1, \mathbf{r}_1) = (1, \mathbf{n}) \cdot (1, \mathbf{r}_2) = 0$$

(i.e.  $\mathbf{n} \cdot \mathbf{r}_1 = \mathbf{n} \cdot \mathbf{r}_2 = -1$ ). It is easy to see that

$$\left( -\frac{x_1y_2 - y_1x_2}{x_1 - x_2}, -\frac{y_1 - y_2}{x_1 - x_2}, 1 \right) \propto (1, \mathbf{n}),$$

and thus the set of solutions can be rewritten as

$$(c_0, \mathbf{c}) = (1, t, 0) + \lambda(1, \mathbf{n}) \quad (87)$$

with  $\lambda \in \mathbb{R}$ . Note that because we are interested in the case when  $A_D^{t\mathbf{x}}$  and  $A_D^{t\mathbf{y}}$  are  $\mathcal{S}_1$ -incompatible, we do not consider the case when  $\mathbf{r}_1$  and  $\mathbf{r}_2$  are parallel or when  $x_1 = x_2$  corresponding to  $\psi_0 = \frac{\pi}{2}$  or  $\varphi_0 = 0$  in Part 1 respectively. Therefore, the vector  $\mathbf{n} = (n_x, n_y)$  can be defined successfully, and it is easy to verify that  $|\mathbf{n}| = \sqrt{n_x^2 + n_y^2} > 1$ . Moreover, because  $\varphi_0$  is supposed to be  $0 < \varphi_0 < \frac{\pi}{2}$  as shown in Part 1, we can assume without loss of generality that its components  $n_x$  and  $n_y$  are negative (see FIG. 4). In order for  $C$  to be an element of  $C(A^{t\mathbf{x}} : \rho^{\mathbf{r}_1}, \rho^{\mathbf{r}_2})$ , (87) should also satisfy

$$0 \leq 1 + \lambda \pm |(t, 0) + \lambda\mathbf{n}| \leq 2,$$

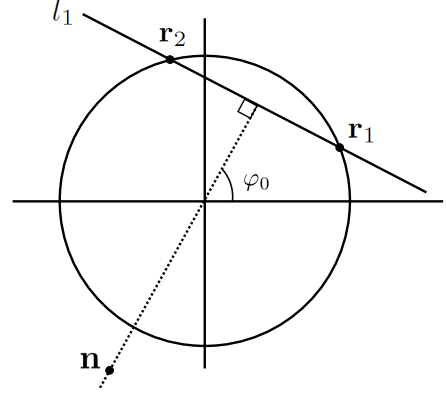


FIG. 4: Geometric description of  $\mathbf{n}$ : we can observe that it lies in the third quadrant.

i.e.

$$1 + \lambda - |(t, 0) + \lambda\mathbf{n}| \geq 0, \quad 1 + \lambda + |(t, 0) + \lambda\mathbf{n}| \leq 2.$$

It can be reduced to

$$\lambda_1^t \leq \lambda \leq \lambda_2^t \quad (88)$$

with

$$\lambda_1^t = \frac{1 - n_x t - \sqrt{(1 - n_x t)^2 + (|\mathbf{n}|^2 - 1)(1 - t^2)}}{|\mathbf{n}|^2 - 1},$$

$$\lambda_2^t =$$

$$\min \left\{ 1, \frac{-1 - n_x t + \sqrt{(1 + n_x t)^2 + (|\mathbf{n}|^2 - 1)(1 - t^2)}}{|\mathbf{n}|^2 - 1} \right\}, \quad (89)$$

where we used  $|\mathbf{n}| > 1$  and  $n_x < 0$  (see FIG. 5). Overall,  $C(A_D^{t\mathbf{x}} : \rho^{\mathbf{r}_1}, \rho^{\mathbf{r}_2})$  is isomorphic to the set parameterized as

$$\{(1, t, 0) + \lambda(1, \mathbf{n}) \mid \lambda_1^t \leq \lambda \leq \lambda_2^t\}, \quad (90)$$

where  $\lambda_1^t$  and  $\lambda_2^t$  are shown in (89). Remark that the same argument can be applied for  $C(A^{t\mathbf{y}} : \rho^{\mathbf{r}_1}, \rho^{\mathbf{r}_2})$ .

$$1 + \lambda + |(t, 0) + \lambda\mathbf{n}| \leq 2$$

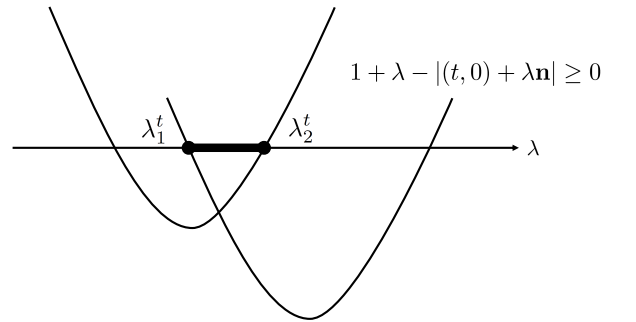


FIG. 5: Solutions for  $\lambda$ .

We shall now prove  $t'_0 = \inf L \notin L$ . Suppose that  $t \in L$ , i.e.  $\chi_{incomp}(A_D^{t\mathbf{x}}, A_D^{t\mathbf{y}}) = 2$ . It follows that there



exist  $\mathbf{r}_1$  and  $\mathbf{r}_2$  in  $\partial D$  such that

$$(C(\mathbf{A}_D^{tx} : \rho^{\mathbf{r}_1}, \rho^{\mathbf{r}_2}) \times C(\mathbf{A}_D^{ty} : \rho^{\mathbf{r}_1}, \rho^{\mathbf{r}_2})) \cap JM(2 : 2) = \emptyset.$$

Denoting  $C(\mathbf{A}_D^{tx} : \rho^{\mathbf{r}_1}, \rho^{\mathbf{r}_2})$  and  $C(\mathbf{A}_D^{ty} : \rho^{\mathbf{r}_1}, \rho^{\mathbf{r}_2})$  simply by  $X^t$  and  $Y^t$  respectively, we can rewrite it as

$$X^t \times Y^t \cap JM(2 : 2) = \emptyset.$$

We need the following lemma.

**Lemma 6.** *Let  $\delta > 0$ . There exists  $\Delta > 0$  such that for all  $\tau \in [0, \Delta]$  and for all  $C \in X^{t-\tau}$ , there exists  $A \in X^t$  satisfying*

$$d(C, A) := \|C(+) - A(+)\| < \delta$$

where  $d$  is a metric on  $O_D(2)$  defined through the operator norm  $\|\cdot\|$  on  $\mathcal{E}_D \simeq O_D(2)$ .

*Proof.* By its definition,  $X^t$  is a convex set of  $O_D(2)$ , and thus for all  $E \in O_D(2)$  we can define successfully the distance between  $E$  and  $X^t$ :

$$d(E, X^t) = \min_{F \in X^t} d(E, F).$$

In particular, for  $E' \in X^{t-\Delta'} \subset O_D(2)$  with  $\Delta' > 0$  and  $E'(+) = \frac{1}{2}(e'_0 \mathbb{1} + \mathbf{e}' \cdot \sigma)$ , it becomes

$$\begin{aligned} d(E', X^t) &= \min_{F \in X^t} d(E', F) \\ &= \min_{F \in X^t} \frac{1}{2} (|e'_0 - f_0| + |\mathbf{e}' - \mathbf{f}|), \end{aligned} \quad (91)$$

where  $F(+) = \frac{1}{2}(f_0 \mathbb{1} + \mathbf{f} \cdot \sigma)$  (see (86)). Since, in terms of (90),  $E' \in X^{t-\Delta'}$  and  $F \in X^t$  imply

$$(e'_0, \mathbf{e}') = (1, t - \Delta', 0) + \lambda'(1, \mathbf{n})$$

with  $\lambda_1^{t-\Delta'} \leq \lambda' \leq \lambda_2^{t-\Delta'}$  and

$$(f_0, \mathbf{f}) = (1, t, 0) + \lambda(1, \mathbf{n})$$

with  $\lambda_1^t \leq \lambda \leq \lambda_2^t$  respectively, (91) can be rewritten as

$$\begin{aligned} 2d(E', X^t) &= \min_{\lambda \in [\lambda_1^t, \lambda_2^t]} (|\lambda' - \lambda| + |(-\Delta', 0) + (\lambda' - \lambda)\mathbf{n}|). \end{aligned}$$

It follows that

$$2d(E', X^t) \leq \Delta' + \min_{\lambda \in [\lambda_1^t, \lambda_2^t]} |\lambda' - \lambda|(1 + |\mathbf{n}|). \quad (92)$$

Let us evaluate its right hand side. It is easy to see that

$$\min_{\lambda \in [\lambda_1^t, \lambda_2^t]} |\lambda' - \lambda| = \begin{cases} \lambda_1^t - \lambda' & (\lambda' < \lambda_1^t) \\ 0 & (\lambda_1^t \leq \lambda' \leq \lambda_2^t) \\ \lambda' - \lambda_2^t & (\lambda' > \lambda_2^t) \end{cases}.$$

Suppose that  $\lambda' < \lambda_1^t$  holds, for example. In this case, because  $\lambda_1^{t-\Delta'} \leq \lambda'$ , we can obtain

$$\lambda_1^t - \lambda' \leq \lambda_1^t - \lambda_1^{t-\Delta'}.$$

In a similar way, it can be demonstrated that

$$\begin{aligned} &\sup_{\lambda' \in [\lambda_1^{t-\Delta'}, \lambda_2^{t-\Delta'}]} \min_{\lambda \in [\lambda_1^t, \lambda_2^t]} |\lambda' - \lambda| \\ &= \max \left\{ \lambda_1^t - \lambda_1^{t-\Delta'}, 0, \lambda_2^{t-\Delta'} - \lambda_2^t \right\}. \end{aligned}$$

By virtue of (89), the right hand side converges to 0 as  $\Delta' \rightarrow 0$ , and thus we can see from (92) that

$$\sup_{E' \in X^{t-\Delta'}} d(E', X^t) \xrightarrow{\Delta' \rightarrow 0} 0$$

It results in that there exists  $\Delta > 0$  such that for all  $\tau \in [0, \Delta]$ ,

$$\sup_{E' \in X^{t-\tau}} d(E', X^t) < \delta$$

holds, that is,  $d(C, X^t) < \delta$  holds for any  $C \in X^{t-\tau}$ . Moreover, because  $X^t$  is convex, there exists  $A \in X^t$  satisfying  $d(C, X^t) = d(C, A)$ , which proves the claim of the lemma.  $\square$

Note that a similar statement also holds for  $Y^t$ : there exists  $\tilde{\Delta} > 0$  such that for all  $\tilde{\tau} \in [0, \tilde{\Delta}]$  and for all  $D \in Y^{t-\tilde{\tau}}$ , there exists  $B \in Y^t$  satisfying  $d(D, B) < \delta$ . Let  $V := O_D(2) \times O_D(2) (\simeq \mathcal{E}_D \times \mathcal{E}_D)$  and let  $d_V$  be a product metric on  $V$  defined as

$$d_V((A, B), (C, D)) = \max\{d(A, C), d(B, D)\}.$$

According to Lemma 6 and its  $Y^t$ -counterpart, if we take  $\Delta_0 = \min\{\Delta, \tilde{\Delta}\} (> 0)$ , then there exists  $(A, B) \in X^t \times Y^t$  for all  $(C, D) \in X^{t-\Delta_0} \times Y^{t-\Delta_0}$  such that  $d_V((A, B), (C, D)) < \delta$ . On the other hand, as we have seen, it holds that

$$X^t \times Y^t \cap JM(2 : 2) = \emptyset.$$

Since  $X^t \times Y^t$  and  $JM(2 : 2)$  are closed in  $V$ , and  $V$  is a metric space, we can apply Urysohn's Lemma [33]. It follows that there exists a continuous (in fact uniformly continuous since  $V$  is compact) function  $f: V \rightarrow [0, 1]$  satisfying  $f(U) = 0$  for any  $U \in X^t \times Y^t$  and  $f(W) = 1$  for any  $W \in JM(2 : 2)$ . The uniform continuity of  $f$  implies that for some  $\varepsilon \in (0, 1)$ , there is  $\delta > 0$  such that

$$\begin{aligned} d_V((E', F'), (E, F)) &< \delta \\ \Rightarrow |f((E', F')) - f((E, F))| &< \varepsilon \end{aligned} \quad (93)$$

holds for any  $(E, F) \in V$ . For this  $\delta$ , we can apply the argument above: we can take  $\Delta_0 > 0$  such that for any  $(C, D) \in X^{t-\Delta_0} \times Y^{t-\Delta_0}$ , there exists  $(A, B) \in X^t \times Y^t$  satisfying  $d_V((A, B), (C, D)) < \delta$ . Because  $f((A, B)) = 0$ , we have  $f((C, D)) < \varepsilon < 1$  (see (93)), and thus  $(C, D) \notin JM(2 : 2)$ . It indicates that  $X^{t-\Delta_0} \times Y^{t-\Delta_0} \cap JM(2 : 2) = \emptyset$ , that is, there is  $\Delta_0 > 0$  for any  $t \in L$  satisfying  $t - \Delta_0 \in L$ . Therefore,  $t'_0 = \inf L \notin L$  can be concluded.

## VI. CONCLUSION

In this study, we have introduced the notions of incompatibility and compatibility dimensions for collections of devices. They describe the minimum number of states which are needed to detect incompatibility and the maximum number of states on which incompatibility vanishes, respectively. We have not only presented general properties of those quantities but also examined concrete behaviors of them for a pair of unbiased qubit observables. We have proved that even for this simple pair of incompatible observables there exist two types of incompatibility with different incompatibility dimensions which cannot be observed if we focus only on robustness of incompatibility under noise. We expect that it is possible to apply this difference to some quantum protocols such as quantum

cryptography. Future work will be needed to investigate whether similar results can be obtained for observables in higher dimensional Hilbert space or other quantum devices. As the definitions apply to devices in GPTs, an interesting task is further to see how quantum incompatibility dimension differ from incompatibility dimension in general.

## Acknowledgments

The authors wish to thank Yui Kuramochi for helpful comments. T.M. acknowledges financial support from JSPS KAKENHI Grant Number JP20K03732. R.T. acknowledges financial support from JSPS KAKENHI Grant Number JP21J10096.

- 
- [1] T. Heinosaari, T. Miyadera, and M. Ziman. An invitation to quantum incompatibility. *J. Phys. A: Math. Theor.*, 49:123001, 2016.
  - [2] C.H. Bennett and G. Brassard. Quantum cryptography: Public-key distribution and coin-tossing. In *Proceedings of IEEE International Conference on Computers, Systems and Signal Processing*, pages 175–179, New York, 1984. IEEE.
  - [3] S. Designolle, M. Farkas, and J. Kaniewski. Incompatibility robustness of quantum measurements: a unified framework. *New J. Phys.*, 21:113053, 2019.
  - [4] T. Heinosaari, J. Schultz, A. Toigo, and M. Ziman. Maximally incompatible quantum observables. *Phys. Lett. A*, 378:1695–1699, 2014.
  - [5] T. Heinosaari, J. Kiukas, D. Reitzner, and J. Schultz. Incompatibility breaking quantum channels. *J. Phys. A: Math. Theor.*, 48:435301, 2015.
  - [6] Y. Kuramochi. Compact convex structure of measurements and its applications to simulability, incompatibility, and convex resource theory of continuous-outcome measurements. arXiv:2002.03504 [math.FA], 2020.
  - [7] M. Plávala. General probabilistic theories: An introduction. arXiv:2103.07469 [quant-ph], 2021.
  - [8] L. Guerini, M.T. Quintino, and L. Aolita. Distributed sampling, quantum communication witnesses, and measurement incompatibility. *Phys. Rev. A*, 100:042308, 2019.
  - [9] J. Kiukas. Subspace constraints for joint measurability. *J. Phys.: Conf. Ser.*, 1638:012003, 2020.
  - [10] F. Loulidi and I. Nechita. The compatibility dimension of quantum measurements. *J. Math. Phys.*, 62:042205, 2021.
  - [11] R. Uola, T. Kraft, S. Designolle, N. Miklin, A. Tavakoli, J.-P. Pellonpää, O. Gühne, and N. Brunner. Quantum measurement incompatibility in subspaces. *Phys. Rev. A*, 103:022203, 2021.
  - [12] T. Heinosaari and M. Ziman. *The Mathematical Language of Quantum Theory*. Cambridge University Press, Cambridge, 2012.
  - [13] P. Busch, T. Heinosaari, J. Schultz, and N. Stevens. Comparing the degrees of incompatibility inherent in probabilistic physical theories. *EPL*, 103:10002, 2013.
  - [14] A. Jenčová and M. Plávala. Conditions on the existence of maximally incompatible two-outcome measurements in general probabilistic theory. *Phys. Rev. A*, 96:022113, 2017.
  - [15] R. Takakura and T. Miyadera. Preparation uncertainty implies measurement uncertainty in a class of generalized probabilistic theories. *J. Math. Phys.*, 61:082203, 2020.
  - [16] P. Busch. Unsharp reality and joint measurements for spin observables. *Phys. Rev. D*, 33:2253–2261, 1986.
  - [17] T. Heinosaari, T. Miyadera, and D. Reitzner. Strongly incompatible quantum devices. *Found. Phys.*, 44:34–57, 2014.
  - [18] E. Haapasalo. Robustness of incompatibility for quantum devices. *J. Phys. A: Math. Theor.*, 48:255303, 2015.
  - [19] T. Heinosaari and T. Miyadera. Incompatibility of quantum channels. *J. Phys. A: Math. Theor.*, 50:135302, 2017.
  - [20] T. Heinosaari, D. Reitzner, T. Rybár, and M. Ziman. Incompatibility of unbiased qubit observables and Pauli channels. *Phys. Rev. A*, 97:022112, 2018.
  - [21] Y. Kuramochi. Quantum incompatibility of channels with general outcome operator algebras. *J. Math. Phys.*, 59:042203, 2018.
  - [22] E. Haapasalo. Compatibility of covariant quantum channels with emphasis on Weyl symmetry. *Ann. Henri Poincaré*, 20:3163, 2019.
  - [23] H. Barnum, J. Barrett, M. Leifer, and A. Wilce. Generalized no-broadcasting theorem. *Phys. Rev. Lett.*, 99:240501, 2007.
  - [24] H. Martens and W.M. de Muynck. Nonideal quantum measurements. *Found. Phys.*, 20:255–281, 1990.
  - [25] H. Barnum, C.M. Caves, C.A. Fuchs, R. Jozsa, and B. Schumacher. Noncommuting mixed states cannot be broadcast. *Phys. Rev. Lett.*, 76:2818–2821, 1996.
  - [26] A. Jenčová. Incompatible measurements in a class of general probabilistic theories. *Phys. Rev. A*, 98:012133, 2018.
  - [27] C. Carmeli, T. Heinosaari, and T. Toigo. Quantum incompatibility witnesses. *Phys. Rev. Lett.*, 122:130402, 2019.
  - [28] C. Carmeli, T. Heinosaari, T. Miyadera, and A. Toigo. Witnessing incompatibility of quantum channels. *J. Math. Phys.*, 60:122202, 2019.

- [29] R. Beneduci, T.J. Bullock, P. Busch, C. Carmeli, T. Heinosaari, and A. Toigo. Operational link between mutually unbiased bases and symmetric informationally complete positive operator-valued measures. *Phys. Rev. A*, 88:032312, 2013.
- [30] P. Stano, D. Reitzner, and T. Heinosaari. Coexistence of qubit effects. *Phys. Rev. A*, 78:012315, 2008.
- [31] P. Busch and H.-J. Schmidt. Coexistence of qubit effects. *Quantum Inf. Process.*, 9:143–169, 2010.
- [32] S. Yu, N.-L. Liu, L. Li, and C.H. Oh. Joint measurement of two unsharp observables of a qubit. *Phys. Rev. A*, 81:062116, 2010.
- [33] J.L. Kelley. *General topology*. Springer-Verlag, New York, 1975. Reprint of the 1955 edition [Van Nostrand, Toronto, Ont.], Graduate Texts in Mathematics, No. 27.