

POST-PROCESSING MINIMAL JOINT OBSERVABLES

TEIKO HEINOSAARI[◇] AND YUI KURAMOCHI[□]

ABSTRACT. A finite set of quantum observables (positive operator valued measures) is called compatible if these observables are marginals of a some observable, called a joint observable of them. For a given set of compatible observables, their joint observable is in general not unique and it is desirable to take a minimal joint observable in the post-processing order since a less informative observable disturbs less the system. We address the question of the minimality of finite-outcome joint observables and prove that any joint observable is lower bounded by a minimal joint observable in the post-processing order. We also give characterizations of the minimality of a joint observable that can be checked by finite-step algorithms and apply them to the case of non-commuting dichotomic qubit observables.

1. INTRODUCTION

An important and intriguing feature of quantum observables is that two observables may not allow a simultaneous measurement, or any other kind of joint implementation. This relation is called incompatibility, and it links interestingly to various other features of quantum theory [1].

Mathematically speaking, quantum observables are described as positive operator valued measures (POVMs). Two observables A and B are compatible if there exists a third observable C that gives both of them as marginals. In this case, C is called their joint observable. The set of all joint observables of a compatible pair of observables is convex. Therefore, a compatible pair has either a unique joint observable or infinitely many different joint observables. If two observables are compatible and at least one of them is sharp, then their joint observable is unique [2]. Other criteria leading to the existence of a unique joint observable have been studied in [3, 4]. These are, in any case, very special

[◇] QTF CENTRE OF EXCELLENCE, DEPARTMENT OF PHYSICS AND ASTRONOMY, UNIVERSITY OF TURKU, TURKU 20014, FINLAND

[□] LABORATORY OF QUANTUM ENGINEERING AND QUANTUM METROLOGY, SCHOOL OF PHYSICS AND ASTRONOMY, SUN YAT-SEN UNIVERSITY (ZHUHAI CAMPUS), ZHUHAI 519082, CHINA

situations and in the generic case there is no unique joint observable. This raises the question on the possible, physically motivated, hierarchy in the set of all joint observables. As a practical problem, one may ask if some joint observables are more preferred than others.

A possible starting point, adopted in the current investigation, is that one should choose a joint observable that allows least disturbing measurement among all joint observables. As explained in Section 2, this means that we are interested in the post-processing ordering of the set joint observables, and seek minimal elements in that set. Our main result is that any joint observable is lower bounded by a minimal joint observable (Section 3, Theorem 1 and Corollary 1). We also give characterizations of minimal joint observables that can be checked by finite-step algorithms for a given joint observable (Section 4). We apply these characterizations to two non-commutative dichotomic qubit observables and obtain a complete characterization of minimality of joint observables in this case (Theorem 3).

2. MOTIVATION OF THE QUESTION

2.1. Post-processing minimal observables. In this work we are going to deal with observables with finite number of outcomes. Let \mathcal{H} be a complex Hilbert space. We denote by $\mathcal{L}(\mathcal{H})$ the set of all bounded operators on \mathcal{H} . An observable is a map $\mathbf{A} : \Omega \rightarrow \mathcal{L}(\mathcal{H})$ such that $\mathbf{A}(x) \geq 0$ for every $x \in \Omega$ and $\sum_x \mathbf{A}(x) = \mathbb{1}$, where Ω is a finite set of measurement outcomes.

As an introductory example, suppose our aim is to discriminate two orthogonal pure states (i.e. one-dimensional projections) $P_1 = |\psi_1\rangle\langle\psi_1|$ and $P_2 = |\psi_2\rangle\langle\psi_2|$ of a, say, 4-dimensional quantum system. One possibility is to complete the set $\{\psi_1, \psi_2\}$ into an orthonormal basis $\{\psi_1, \psi_2, \psi_3, \psi_4\}$ and then make a measurement in that basis. The corresponding observable is hence $\mathbf{A}(x) = |\psi_x\rangle\langle\psi_x|$, $x = 1, \dots, 4$. This kind of measurement, however, disturbs the system more than is necessary. Instead of measuring \mathbf{A} , we can, for instance, perform a measurement of a two-outcome observable \mathbf{B} , defined as $\mathbf{B}(1) = |\psi_1\rangle\langle\psi_1| + |\psi_3\rangle\langle\psi_3|$, $\mathbf{B}(2) = |\psi_2\rangle\langle\psi_2| + |\psi_4\rangle\langle\psi_4|$. The observable \mathbf{B} discriminates the states P_1 and P_2 , but allows a less disturbing measurement than \mathbf{A} . In this exemplary case, we can get \mathbf{B} from \mathbf{A} by grouping the outcomes. Namely, we have

$$\mathbf{B}(1) = \mathbf{A}(1) + \mathbf{A}(3), \quad \mathbf{B}(2) = \mathbf{A}(2) + \mathbf{A}(4).$$

This sort of relabeling of outcomes is a special type of post-processing.

To recall the general definition of post-processing, let Ω_1 and Ω_2 be finite sets. A map $p: \Omega_1 \times \Omega_2 \rightarrow \mathbb{R}$ is called a *Markov kernel*, or

stochastic matrix, from Ω_2 to Ω_1 if for each $x \in \Omega_1$ and each $y \in \Omega_2$, $p(x, y) \geq 0$ and $\sum_{x' \in \Omega_1} p(x', y) = 1$. The set of Markov kernels from Ω_2 to Ω_1 is written as $\mathbf{Markov}(\Omega_1, \Omega_2)$, and it is a compact convex subset of the Euclidean space $\mathbb{R}^{\Omega_1 \times \Omega_2}$. Let $\mathbf{A}: \Omega_1 \rightarrow \mathcal{L}(\mathcal{H})$ be an observable, let Ω_2 be a finite set, and let $p \in \mathbf{Markov}(\Omega_2, \Omega_1)$. We define an observable $p * \mathbf{A}: \Omega_2 \rightarrow \mathcal{L}(\mathcal{H})$ by

$$p * \mathbf{A}(x) := \sum_{y \in \Omega_1} p(x, y) \mathbf{A}(y).$$

If an observable \mathbf{B} , with an outcome set Ω_2 , can be written as $p * \mathbf{A}$ for some $p \in \mathbf{Markov}(\Omega_2, \Omega_1)$, then \mathbf{B} is called a *post-processing* of \mathbf{A} , and written as $\mathbf{B} \preceq_{\text{post}} \mathbf{A}$. This relation has been studied e.g. in [5, 6, 7].

The relation \preceq_{post} is reflexive and transitive, hence a preorder. A preorder induces an equivalence relation \sim_{post} and a partial order in the set of equivalence classes. We write $\mathbf{A} \prec_{\text{post}} \mathbf{B}$ if $\mathbf{A} \preceq_{\text{post}} \mathbf{B}$ holds but $\mathbf{B} \preceq_{\text{post}} \mathbf{A}$ does not hold. We will say that an observable \mathbf{A} is maximal/minimal/greatest/least in a subset X if the corresponding equivalence class $[\mathbf{A}]$ has that property in the respective subset of equivalence classes. For instance, we say that an observable \mathbf{A} is *post-processing minimal (in X)* if the following implication holds for every observable $\mathbf{B} (\in X)$:

$$\mathbf{B} \preceq_{\text{post}} \mathbf{A} \quad \Rightarrow \quad \mathbf{A} \preceq_{\text{post}} \mathbf{B}.$$

It is known that an observable \mathbf{A} is post-processing minimal in the set of all observables if and only if each of its operator is a multiple of the identity operator $\mathbb{1}$, while \mathbf{A} is post-processing maximal in the set of all observables if and only if each of its nonzero operator is rank-1 [5].

It has been shown in [8] that for two observables \mathbf{A} and \mathbf{B} , the relation $\mathbf{B} \preceq_{\text{post}} \mathbf{A}$ holds if and only if the disturbance related to \mathbf{B} is smaller than or equal to the disturbance related to \mathbf{A} . The disturbance related to an observable \mathbf{A} refers to the set of all quantum channels that arise in some measurement of \mathbf{A} . As we may want to perform subsequent measurements, it is desirable to disturb the initial state as little as possible. We thereby take the following as a guiding principle:

Whenever we need to choose an observable that has certain property (e.g. enables discrimination of some states), we should choose it to be minimal in the post-processing preorder in the set of all observables with that property.

In a typical case, this guiding principle does not lead to a unique choice. For instance, in the starting example we can also choose $\mathbf{B}'(1) =$

$|\psi_1\rangle\langle\psi_1| + |\psi_4\rangle\langle\psi_4|$, $\mathbf{B}'(2) = |\psi_2\rangle\langle\psi_2| + |\psi_3\rangle\langle\psi_3|$, or any convex combination of \mathbf{B} and \mathbf{B}' . In this specific example any two-outcome discriminating observable is post-processing minimal. To see this, we first note that a two-outcome observable \mathbf{C} discriminates ψ_1 and ψ_2 if and only if $\mathbf{C}(i) \geq P_i$, $i = 1, 2$. Let \mathbf{C} and \mathbf{C}' be two-outcome observables that discriminate ψ_1 and ψ_2 and suppose that $\mathbf{C} \preceq_{\text{post}} \mathbf{C}'$. We take a Markov kernel p such that $\mathbf{C} = p * \mathbf{C}'$. Then

$$P_1 = P_1\mathbf{C}(1) = \sum_{i=1,2} p(1,i)P_1\mathbf{C}'(i) = p(1,1)P_1$$

$$P_2 = P_2\mathbf{C}(2) = \sum_{i=1,2} p(2,i)P_2\mathbf{C}'(i) = p(2,2)P_2$$

Therefore $p(1,1) = p(2,2) = 1$ and hence $p(1,2) = p(2,1) = 0$. Thus $\mathbf{C} = \mathbf{C}'$. This implies the post-processing minimality of any two-outcome discriminating observable.

2.2. Minimal joint observables. In our investigation we consider a situation where there are several tasks that we want to perform. We assume that observables $\mathbf{A}_1, \dots, \mathbf{A}_n$ for the individual tasks have already been chosen and that they are compatible. We recall that observables $\mathbf{A}_1, \dots, \mathbf{A}_n$ are, by definition, compatible if there exists an observable \mathbf{C} such that $\mathbf{A}_\ell \preceq_{\text{post}} \mathbf{C}$ for all $\ell = 1, \dots, n$. Clearly, measuring \mathbf{C} is then enough to simulate measurements of all $\mathbf{A}_1, \dots, \mathbf{A}_n$.

According to our guiding principle, we want to choose a post-processing minimal observable among all observables \mathbf{C} that satisfy $\mathbf{A}_\ell \preceq_{\text{post}} \mathbf{C}$ for all $\ell = 1, \dots, n$. Our first observation, based on [9], is that if such \mathbf{C} exists, then there also exists an observable \mathbf{G} with the outcome set $\Omega_1 \times \dots \times \Omega_n$ such that $\mathbf{G} \preceq_{\text{post}} \mathbf{C}$ and each \mathbf{A}_ℓ is the ℓ th marginal of \mathbf{G} , i.e.,

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

$$\vdots$$

$$\sum_{x_1, \dots, x_{n-1}} \mathbf{G}(x_1, x_2, \dots, x_n) = \mathbf{A}_n(x_n)$$

This kind of observable is called a *joint observable* of $\mathbf{A}_1, \dots, \mathbf{A}_n$ [10].

To verify the previous claims, we first recall that $A_\ell \preceq_{\text{post}} C$ means that there exists a Markov kernel p_ℓ such that

$$A_\ell(x) = \sum_y p_\ell(x, y) C(y).$$

We define

$$p((x_1, \dots, x_n), y) := \prod_{\ell=1}^n p_\ell(x_\ell, y) \quad (1)$$

and set

$$G(x_1, \dots, x_n) := \sum_y p((x_1, \dots, x_n), y) C(y). \quad (2)$$

It is straightforward to verify that G is a joint observable of A_1, \dots, A_n . Further, p is a Markov kernel and, thus, (2) means that $G \preceq_{\text{post}} C$.

We conclude that when we search for minimal observables C satisfying $A_\ell \preceq_{\text{post}} C$ for all $\ell = 1, \dots, n$, we can limit our search for joint observables of A_1, \dots, A_n . Namely, any C with the required property is either post-processing equivalent to some joint observable, or strictly greater than some joint observable.

For a finite set of observables $\{A_\ell\}_{\ell=1}^n$, we denote by $\mathcal{J}(\{A_\ell\}_{\ell=1}^n)$ the set of all their joint observables. The following concept will be our main focus.

Definition 1. Let $\{A_\ell\}_{\ell=1}^n$ be a set of compatible observables. Their joint observable G is a *minimal joint observable* if G is post-processing minimal in $\mathcal{J}(\{A_\ell\}_{\ell=1}^n)$. The set of minimal observables of $\{A_\ell\}_{\ell=1}^n$ is denoted by $\mathcal{J}_{\min}(\{A_\ell\}_{\ell=1}^n)$.

Since the post-processing relation is a preorder rather than a partial order, it is often convenient to work with the equivalence classes of observables, and in that case the induced post-processing relation is a partial order. We denote by $\mathcal{J}_{\min}(\{A_\ell\}_{\ell=1}^n) / \sim_{\text{post}}$ the partially ordered set of equivalence classes of minimal joint observables.

3. ORDER STRUCTURE OF THE SET OF JOINT OBSERVABLES

In this section $\{A_\ell\}_{\ell=1}^n$ is a fixed set of compatible observables. The condition of minimality for a joint observable is independent of the choice of the post-processing representatives of the marginal observables as shown in the following proposition.

Proposition 1. Let $\{A_\ell\}_{\ell=1}^n$ be a set of compatible observables and let $\{A'_\ell\}_{\ell=1}^n$ be observables such that $A_\ell \sim_{\text{post}} A'_\ell$ for all ℓ . Then there exists a bijection

$$f: \mathcal{J}_{\min}(\{A_\ell\}_{\ell=1}^n) / \sim_{\text{post}} \rightarrow \mathcal{J}_{\min}(\{A'_\ell\}_{\ell=1}^n) / \sim_{\text{post}}$$

such that $[G] \sim_{\text{post}} f([G])$ for all $G \in \mathcal{J}_{\min}(\{A_\ell\}_{\ell=1}^n)$.

Proof. Let $G \in \mathcal{J}_{\min}(\{A_\ell\}_{\ell=1}^n)$. Since $A'_\ell \sim_{\text{post}} A_\ell \preceq_{\text{post}} G$, there exists a joint observable $G' \in \mathcal{J}(\{A'_\ell\}_{\ell=1}^n)$ such that $G' \preceq_{\text{post}} G$. Take an arbitrary $G'_0 \in \mathcal{J}(\{A'_\ell\}_{\ell=1}^n)$ satisfying $G'_0 \preceq_{\text{post}} G'$. Since $A_\ell \preceq_{\text{post}} G'_0$ for all ℓ , there exists a joint observable $G_0 \in \mathcal{J}(\{A_\ell\}_{\ell=1}^n)$ such that $G_0 \preceq_{\text{post}} G'_0$. Then we have $G_0 \preceq_{\text{post}} G'_0 \preceq_{\text{post}} G' \preceq_{\text{post}} G$ and the minimality of G implies that all of the joint observables G_0 , G'_0 and G' are post-processing equivalent to G . This shows that G' is minimal and post-processing equivalent to G . Hence there exists a mapping $f: \mathcal{J}_{\min}(\{A_\ell\}_{\ell=1}^n)/\sim_{\text{post}} \rightarrow \mathcal{J}_{\min}(\{A'_\ell\}_{\ell=1}^n)/\sim_{\text{post}}$ such that $[G] \sim_{\text{post}} f([G])$ for all $G \in \mathcal{J}_{\min}(\{A_\ell\}_{\ell=1}^n)$. f is apparently injective. By interchanging $\{A_\ell\}_{\ell=1}^n$ with $\{A'_\ell\}_{\ell=1}^n$ in the above discussion, we can also conclude the existence of an injection $g: \mathcal{J}_{\min}(\{A'_\ell\}_{\ell=1}^n)/\sim_{\text{post}} \rightarrow \mathcal{J}_{\min}(\{A_\ell\}_{\ell=1}^n)/\sim_{\text{post}}$ such that $[G'] \sim_{\text{post}} g([G'])$ for all $G' \in \mathcal{J}_{\min}(\{A'_\ell\}_{\ell=1}^n)$. Then g is the inverse map of f and hence f is a bijection. \square

The following theorem states that any post-processing monotone net of joint observables has its supremum or infimum that is also a joint observable.

Theorem 1. *Let $\{A_\ell\}_{\ell=1}^n$ be a finite set of compatible observables and let $([G_\lambda])_{\lambda \in \Lambda}$ be a net in $\mathcal{J}(\{A_\ell\}_{\ell=1}^n)/\sim_{\text{post}}$.*

1. *If $([G_\lambda])_{\lambda \in \Lambda}$ is monotonically increasing, then there exists a joint observable $\tilde{G} \in \mathcal{J}(\{A_\ell\}_{\ell=1}^n)$ such that $[\tilde{G}]$ is an upper bound of the net $([G_\lambda])_{\lambda \in \Lambda}$ and for any observable B , $G_\lambda \preceq_{\text{post}} B$ ($\forall \lambda \in \Lambda$) implies $\tilde{G} \preceq_{\text{post}} B$.*
2. *If $([G_\lambda])_{\lambda \in \Lambda}$ is monotonically decreasing, then there exists a joint observable $\tilde{G} \in \mathcal{J}(\{A_\ell\}_{\ell=1}^n)$ such that $[\tilde{G}]$ is a lower bound of the net $([G_\lambda])_{\lambda \in \Lambda}$ and for any observable B , $B \preceq_{\text{post}} G_\lambda$ ($\forall \lambda \in \Lambda$) implies $B \preceq_{\text{post}} \tilde{G}$.*

Proof. Let Ω_ℓ be the outcome set of A_ℓ and let $\tilde{\Omega} := \Omega_1 \times \cdots \times \Omega_n$.

1. By assumption, for $\lambda, \lambda' \in \Lambda$ satisfying $\lambda \leq \lambda'$ there exists $p^{\lambda|\lambda'} \in \mathbf{Markov}(\tilde{\Omega}, \tilde{\Omega})$ such that $G_\lambda = p^{\lambda|\lambda'} * G_{\lambda'}$. For each $\lambda, \lambda' \in \Lambda$ we define a Markov kernel $q^{\lambda|\lambda'} \in \mathbf{Markov}(\tilde{\Omega}, \tilde{\Omega})$ by

$$q^{\lambda|\lambda'}(\mathbf{x}, \mathbf{x}') := \begin{cases} p^{\lambda|\lambda'}(\mathbf{x}, \mathbf{x}') & \text{if } \lambda \leq \lambda', \\ \delta_{\mathbf{x}, \mathbf{x}'} & \text{otherwise.} \end{cases}$$

By Tychonoff's theorem, the set $\mathbf{Markov}(\tilde{\Omega}, \tilde{\Omega})^\Lambda \times \mathcal{J}(\{A_\ell\}_{\ell=1}^n)$ is compact in the product topology. Hence there exists a subnet $(G_{\lambda'(i)})_{i \in I}$ such that $q^{\lambda|\lambda'(i)}$ converges to some $q^\lambda \in \mathbf{Markov}(\tilde{\Omega}, \tilde{\Omega})$ for each

$\lambda \in \Lambda$ and $\mathbf{G}_{\lambda'(i)}$ converges to some $\tilde{\mathbf{G}} \in \mathcal{J}(\{\mathbf{A}_\ell\}_{\ell=1}^n)$. For each $\lambda \in \Lambda$ and each $i \in I$ satisfying $\lambda \leq \lambda'(i)$, we have $\mathbf{G}_\lambda = p^{\lambda|\lambda'(i)} * \mathbf{G}_{\lambda'(i)} = q^{\lambda|\lambda'(i)} * \mathbf{G}_{\lambda'(i)}$. Thus

$$\begin{aligned} & \left\| \mathbf{G}_\lambda(\mathbf{x}) - \sum_{\mathbf{x}' \in \tilde{\Omega}} q^\lambda(\mathbf{x}, \mathbf{x}') \tilde{\mathbf{G}}(\mathbf{x}') \right\| \\ &= \left\| \sum_{\mathbf{x}' \in \tilde{\Omega}} q^{\lambda|\lambda'(i)}(\mathbf{x}, \mathbf{x}') \mathbf{G}_{\lambda'(i)}(\mathbf{x}') - \sum_{\mathbf{x}' \in \tilde{\Omega}} q^\lambda(\mathbf{x}, \mathbf{x}') \tilde{\mathbf{G}}(\mathbf{x}') \right\| \\ &\leq \sum_{\mathbf{x}' \in \tilde{\Omega}} |q^{\lambda|\lambda'(i)}(\mathbf{x}, \mathbf{x}') - q^\lambda(\mathbf{x}, \mathbf{x}')| \|\mathbf{G}_{\lambda'(i)}(\mathbf{x}')\| + \sum_{\mathbf{x}' \in \tilde{\Omega}} q^\lambda(\mathbf{x}, \mathbf{x}') \|\mathbf{G}_{\lambda'(i)}(\mathbf{x}') - \tilde{\mathbf{G}}(\mathbf{x}')\| \\ &\leq \sum_{\mathbf{x}' \in \tilde{\Omega}} |q^{\lambda|\lambda'(i)}(\mathbf{x}, \mathbf{x}') - q^\lambda(\mathbf{x}, \mathbf{x}')| + \sum_{\mathbf{x}' \in \tilde{\Omega}} q^\lambda(\mathbf{x}, \mathbf{x}') \|\mathbf{G}_{\lambda'(i)}(\mathbf{x}') - \tilde{\mathbf{G}}(\mathbf{x}')\| \\ &\rightarrow 0, \end{aligned}$$

where $\|\cdot\|$ denotes the operator norm. This implies $\mathbf{G}_\lambda = q^\lambda * \tilde{\mathbf{G}} \preceq_{\text{post}} \tilde{\mathbf{G}}$. Hence $[\tilde{\mathbf{G}}]$ is an upper bound of the net $([\mathbf{G}_\lambda])_{\lambda \in \Lambda}$.

Now we take an observable $(\mathbf{B}(y))_{y \in \Omega'}$ satisfying $\mathbf{G}_\lambda \preceq_{\text{post}} \mathbf{B}$ for all $\lambda \in \Lambda$. Then for each $\lambda \in \Lambda$ there exists $r^\lambda \in \mathbf{Markov}(\tilde{\Omega}, \Omega')$ such that $\mathbf{G}_\lambda = r^\lambda * \mathbf{B}$. Since $\mathbf{Markov}(\tilde{\Omega}, \Omega')$ is compact, there exists a subnet $(r^{\lambda''(j)})_{j \in J}$ of $(r^{\lambda'(i)})_{i \in I}$ such that $r^{\lambda''(j)}$ converges to some $\tilde{r} \in \mathbf{Markov}(\tilde{\Omega}, \Omega')$. Thus

$$\tilde{\mathbf{G}}(\mathbf{x}) = \lim_{j \in J} \mathbf{G}_{\lambda''(j)}(\mathbf{x}) = \lim_{j \in J} \sum_{y \in \Omega'} r^{\lambda''(j)}(\mathbf{x}, y) \mathbf{B}(y) = \sum_{y \in \Omega'} \tilde{r}(\mathbf{x}, y) \mathbf{B}(y),$$

which implies $\tilde{\mathbf{G}} = \tilde{r} * \mathbf{B} \preceq_{\text{post}} \mathbf{B}$.

2. For each joint observable $\mathbf{G} \in \mathcal{J}(\{\mathbf{A}_\ell\}_{\ell=1}^n)$ we define

$$K_{\mathbf{G}} := \left\{ p \in \mathbf{Markov}(\tilde{\Omega}, \tilde{\Omega}) \mid p * \mathbf{G} \in \mathcal{J}(\{\mathbf{A}_\ell\}_{\ell=1}^n) \right\},$$

which is a nonempty compact subset of $\mathbf{Markov}(\tilde{\Omega}, \tilde{\Omega})$. By assumption, for $\lambda' \geq \lambda$ there exists a Markov kernel $p^{\lambda'|\lambda} \in K_{\mathbf{G}_\lambda}$ such that $\mathbf{G}_{\lambda'} = p^{\lambda'|\lambda} * \mathbf{G}_\lambda$. For each $\lambda, \lambda' \in \Lambda$, we define $q^{\lambda'|\lambda} \in K_{\mathbf{G}_\lambda}$ by

$$q^{\lambda'|\lambda}(\mathbf{x}', \mathbf{x}) := \begin{cases} p^{\lambda'|\lambda}(\mathbf{x}', \mathbf{x}) & \text{if } \lambda \leq \lambda', \\ \delta_{\mathbf{x}', \mathbf{x}} & \text{otherwise.} \end{cases}$$

By Tychonoff's theorem, the set $\prod_{\lambda \in \Lambda} K_{\mathbf{G}_\lambda}$ is compact in the product topology, and hence there exists a subnet $(\mathbf{G}_{\lambda'(i)})_{i \in I}$ such that

$(q^{\lambda'(i)|\lambda})_{i \in I}$ converges to some $q^\lambda \in K_{\mathbf{G}_\lambda}$ for every $\lambda \in \Lambda$. For each $\lambda \in \Lambda$ and each $i \in I$ satisfying $\lambda'(i) \geq \lambda$, we have

$$\begin{aligned} \mathbf{G}_{\lambda'(i)}(\mathbf{x}') &= \sum_{\mathbf{x} \in \tilde{\Omega}} p^{\lambda'(i)|\lambda}(\mathbf{x}', \mathbf{x}) \mathbf{G}_\lambda(\mathbf{x}) = \sum_{\mathbf{x} \in \tilde{\Omega}} q^{\lambda'(i)|\lambda}(\mathbf{x}', \mathbf{x}) \mathbf{G}_\lambda(\mathbf{x}) \\ &\rightarrow \sum_{\mathbf{x} \in \tilde{\Omega}} q^\lambda(\mathbf{x}', \mathbf{x}) \mathbf{G}_\lambda(\mathbf{x}). \end{aligned}$$

Hence we may define $\tilde{\mathbf{G}} \in \mathcal{J}(\{\mathbf{A}_\ell\}_{\ell=1}^m)$ by

$$\tilde{\mathbf{G}}(\mathbf{x}') := \lim_{i \in I} \mathbf{G}_{\lambda'(i)}(\mathbf{x}') = \sum_{\mathbf{x} \in \tilde{\Omega}} q^\lambda(\mathbf{x}', \mathbf{x}) \mathbf{G}_\lambda(\mathbf{x})$$

By the last equality $[\tilde{\mathbf{G}}]$ is a lower bound of $([\mathbf{G}_\lambda])_{\lambda \in \Lambda}$.

Now we take an observable $(\mathbf{B}(y))_{y \in \Omega'}$ satisfying $\mathbf{B} \preceq_{\text{post}} \mathbf{G}_\lambda$ for all $\lambda \in \Lambda$. Then for each $\lambda \in \Lambda$ there exists a Markov kernel $r^\lambda \in \mathbf{Markov}(\Omega', \tilde{\Omega})$ such that $\mathbf{B} = r^\lambda * \mathbf{G}_\lambda$. Since $\mathbf{Markov}(\Omega', \tilde{\Omega})$ is compact, there exists a subnet $(r^{\lambda''(j)})_{j \in J}$ of $(r^{\lambda'(i)})_{i \in I}$ such that $r^{\lambda''(j)}$ converges to some $\tilde{r} \in \mathbf{Markov}(\Omega', \tilde{\Omega})$. Then we have

$$\begin{aligned} &\left\| \sum_{\mathbf{x} \in \tilde{\Omega}} \tilde{r}(y, \mathbf{x}) \tilde{\mathbf{G}}(\mathbf{x}) - \mathbf{B}(y) \right\| \\ &= \left\| \sum_{\mathbf{x} \in \tilde{\Omega}} \tilde{r}(y, \mathbf{x}) \tilde{\mathbf{G}}(\mathbf{x}) - \sum_{\mathbf{x} \in \tilde{\Omega}} r^{\lambda''(j)}(y, \mathbf{x}) \mathbf{G}_{\lambda''(j)}(\mathbf{x}) \right\| \\ &\leq \sum_{\mathbf{x} \in \tilde{\Omega}} \tilde{r}(y, \mathbf{x}) \left\| \tilde{\mathbf{G}}(\mathbf{x}) - \mathbf{G}_{\lambda''(j)}(\mathbf{x}) \right\| + \sum_{\mathbf{x} \in \tilde{\Omega}} |\tilde{r}(y, \mathbf{x}) - r^{\lambda''(j)}(y, \mathbf{x})| \left\| \mathbf{G}_{\lambda''(j)}(\mathbf{x}) \right\| \\ &\leq \sum_{\mathbf{x} \in \tilde{\Omega}} \tilde{r}(y, \mathbf{x}) \left\| \tilde{\mathbf{G}}(\mathbf{x}) - \mathbf{G}_{\lambda''(j)}(\mathbf{x}) \right\| + \sum_{\mathbf{x} \in \tilde{\Omega}} |\tilde{r}(y, \mathbf{x}) - r^{\lambda''(j)}(y, \mathbf{x})| \\ &\rightarrow 0. \end{aligned}$$

This implies $\mathbf{B} = \tilde{r} * \tilde{\mathbf{G}} \preceq_{\text{post}} \tilde{\mathbf{G}}$, which completes the proof. \square

We can easily see that $[\tilde{\mathbf{G}}]$ in Theorem 1 is a supremum or an infimum of $([\mathbf{G}_\lambda])_{\lambda \in \Lambda}$ in $\mathcal{J}(\{\mathbf{A}_\ell\}_{\ell=1}^n) / \sim_{\text{post}}$. Thus we obtain

Corollary 1. *Let $\{\mathbf{A}_\ell\}_{\ell=1}^n$ be a finite set of compatible observables.*

1. *The poset $\mathcal{J}(\{\mathbf{A}_\ell\}_{\ell=1}^n) / \sim_{\text{post}}$ is both upper and lower directed complete (see Appendix A).*
2. *Any joint observable $\mathbf{G} \in \mathcal{J}(\{\mathbf{A}_\ell\}_{\ell=1}^n)$ is upper bounded by a maximal joint observable and lower bounded by a minimal joint observable.*

Unlike the minimal joint observables, the set of maximal joint observables does depend on the choice of the post-processing representatives of the marginal observables as the following example demonstrates.

Example 1. Let $A_1 = A_2 = (\mathbb{1}_{\mathcal{H}})$ be the single-outcome trivial observable. Then A_1 and A_2 are compatible and the set $\mathcal{J}(\{A_1, A_2\})$ consists of only one element, the trivial observable $(\mathbb{1}_{\mathcal{H}})$, which is both minimal and maximal in $\mathcal{J}(\{A_1, A_2\})$.

Clearly, the observables $A_1 = A_2$ are post-processing equivalent to the two-outcome trivial observable $A'_1 = A'_2 = (\frac{1}{2}\mathbb{1}_{\mathcal{H}}, \frac{1}{2}\mathbb{1}_{\mathcal{H}})$. If $\dim \mathcal{H} \geq 2$, there exists a non-trivial joint observable $G \in \mathcal{J}(\{A'_1, A'_2\})$. Namely, fix an effect $0 \neq T \neq \mathbb{1}$ and define

$$G(1, 1) = G(2, 2) = \frac{1}{2}T, \quad G(1, 2) = G(2, 1) = \frac{1}{2}(\mathbb{1} - T).$$

This is a joint observable of A'_1 and A'_2 . From Corollary 1 we conclude that G is upper bounded by a maximal joint observable $G' \in \mathcal{J}(\{A'_1, A'_2\})$, which is also non-trivial. Hence there is no one-to-one correspondence between $\mathcal{J}_{\max}(\{A_1, A_2\})$ and $\mathcal{J}_{\max}(\{A'_1, A'_2\})$ as in Proposition 1.

The motivation for searching minimal joint observables was explained in Sec. 2.1. We are not aware of a similar motivation for studying maximal joint observables. Let us also note that among the observables C that satisfy $A_\ell \preceq_{\text{post}} C$ for all $\ell = 1, \dots, n$, there is always an observable that is post-processing maximal among the set of all observables, i.e., it consists of rank-1 operators. For these reasons, we focus on minimal joint observables, although we will make some remarks about maximal joint observables as well.

4. CHARACTERIZATION OF MINIMAL JOINT OBSERVABLE

In this section we give an algorithm to determine whether a given joint observable is minimal or not. Throughout this section, we fix a finite set of compatible observables $\{A_\ell\}_{\ell=1}^n$ and a joint observable $G \in \mathcal{J}(\{A_\ell\}_{\ell=1}^n)$. Let Ω_ℓ be the outcome set of A_ℓ and denote $\tilde{\Omega} := \Omega_1 \times \dots \times \Omega_n$.

We need a bunch of auxiliary definitions before we can state our results. For any two observables A and B , with outcome sets Ω and Ω' , respectively, we define

$$K(A, B) := \{p \in \mathbf{Markov}(\Omega, \Omega') \mid A = p * B\},$$

which is, if nonempty, a compact convex subset of $\mathbf{Markov}(\Omega, \Omega')$. Clearly, $K(A, B)$ is nonempty if and only if $A \preceq_{\text{post}} B$.

We define

$$K_{\mathbf{G}} := \{p \in \mathbf{Markov}(\tilde{\Omega}, \tilde{\Omega}) \mid p * \mathbf{G} \in \mathcal{J}(\{\mathbf{A}_\ell\}_{\ell=1}^n)\}.$$

This is the set of those post-processings that are allowed for \mathbf{G} so that it is still a joint observable. For $p \in \mathbb{R}^{\tilde{\Omega} \times \tilde{\Omega}}$, we have $p \in K_{\mathbf{G}}$ if and only if p satisfies the following linear (in)equalities:

$$\begin{aligned} \sum_{\mathbf{x}' \in \tilde{\Omega}} p(\mathbf{x}', \mathbf{x}) &= 1 \quad (\mathbf{x} \in \tilde{\Omega}), \\ \sum_{\mathbf{x}' \in \tilde{\Omega}, \pi_\ell(\mathbf{x}') = x'_\ell} \sum_{\mathbf{x} \in \tilde{\Omega}} p(\mathbf{x}', \mathbf{x}) \mathbf{G}(\mathbf{x}) &= \mathbf{A}_\ell(x'_\ell) \quad (\ell \in \{1, \dots, n\}; x'_\ell \in \Omega_\ell), \\ p(\mathbf{x}', \mathbf{x}) &\geq 0 \quad (\mathbf{x}, \mathbf{x}' \in \tilde{\Omega}), \end{aligned}$$

where $\pi_\ell: \tilde{\Omega} \rightarrow \Omega_\ell$ is the canonical projection. Thus, $K_{\mathbf{G}}$ is a compact convex polytope on the Euclidean space $\mathbb{R}^{\tilde{\Omega} \times \tilde{\Omega}}$. It follows that the set of extreme points of $K_{\mathbf{G}}$ is finite and can be written as $\{p_1, \dots, p_N\}$ (see Appendix C). We define

$$p_* := \frac{1}{N} \sum_{k=1}^N p_k \in K_{\mathbf{G}}.$$

For each $\ell \in \{1, \dots, n\}$, $K(\mathbf{A}_\ell, \mathbf{G})$ is also a nonempty compact convex polytope on $\mathbb{R}^{\Omega_\ell \times \tilde{\Omega}}$ and the set of extreme points of $K(\mathbf{A}_\ell, \mathbf{G})$ can be written as $\{q_{\ell,1}, \dots, q_{\ell,N_\ell}\}$. We define $q_\ell \in K(\mathbf{A}_\ell, \mathbf{G})$ by

$$q_\ell := \frac{1}{N_\ell} \sum_{k=1}^{N_\ell} q_{\ell,k}$$

and $q_* \in K_{\mathbf{G}}$ by

$$q_*(\mathbf{x}', \mathbf{x}) := \prod_{\ell=1}^n q_\ell(\pi_\ell(\mathbf{x}'), \mathbf{x}).$$

Finally, for $r \in \mathbf{Markov}(\tilde{\Omega}, \tilde{\Omega})$, r is said to be *conditionally independent* if there exist Markov kernels $r_\ell \in \mathbf{Markov}(\Omega_\ell, \tilde{\Omega})$ ($\ell \in \{1, \dots, n\}$) such that

$$r(\mathbf{x}', \mathbf{x}) = \prod_{\ell=1}^n r_\ell(\pi_\ell(\mathbf{x}'), \mathbf{x}). \quad (3)$$

We define

$$K_{\mathbf{G}}^{\text{ind}} := \{r \in K_{\mathbf{G}} \mid r \text{ is conditionally independent}\}.$$

Now we can state and prove the following result.

Theorem 2. *The following conditions are equivalent.*

- (i) \mathbf{G} is a minimal joint observable.
- (ii) For each $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}' \in \tilde{\Omega}$ and each $p \in K_{\mathbf{G}}$, if $\mathbf{G}(\mathbf{x}_1)$ and $\mathbf{G}(\mathbf{x}_2)$ are linearly independent then $p(\mathbf{x}', \mathbf{x}_1)p(\mathbf{x}', \mathbf{x}_2) = 0$.
- (iii) For each $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}' \in \tilde{\Omega}$, if $\mathbf{G}(\mathbf{x}_1)$ and $\mathbf{G}(\mathbf{x}_2)$ are linearly independent then $p_*(\mathbf{x}', \mathbf{x}_1)p_*(\mathbf{x}', \mathbf{x}_2) = 0$.
- (iv) $p_* * \mathbf{G} \sim_{\text{post}} \mathbf{G}$.
- (v) $r * \mathbf{G} \sim_{\text{post}} \mathbf{G}$ for all $r \in K_{\mathbf{G}}^{\text{ind}}$.
- (vi) For each $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}' \in \tilde{\Omega}$ and each $(r_\ell)_{\ell=1}^n \in K(\mathbf{A}_\ell, \mathbf{G})$, if $\mathbf{G}(\mathbf{x}_1)$ and $\mathbf{G}(\mathbf{x}_2)$ are linearly independent then there exists $\ell \in \{1, \dots, n\}$ such that $r_\ell(\pi_\ell(\mathbf{x}', \mathbf{x}_1))r_\ell(\pi_\ell(\mathbf{x}', \mathbf{x}_2)) = 0$.
- (vii) For each $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}' \in \tilde{\Omega}$, if $\mathbf{G}(\mathbf{x}_1)$ and $\mathbf{G}(\mathbf{x}_2)$ are linearly independent then there exists $\ell \in \{1, \dots, n\}$ such that $p(\pi_\ell(\mathbf{x}', \mathbf{x}_1))p(\pi_\ell(\mathbf{x}', \mathbf{x}_2)) = 0$ for all $p \in K(\mathbf{A}_\ell, \mathbf{G})$.
- (viii) For each $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}' \in \tilde{\Omega}$, if $\mathbf{G}(\mathbf{x}_1)$ and $\mathbf{G}(\mathbf{x}_2)$ are linearly independent then there exists $\ell \in \{1, \dots, n\}$ such that $q_\ell(\pi_\ell(\mathbf{x}', \mathbf{x}_1))q_\ell(\pi_\ell(\mathbf{x}', \mathbf{x}_2)) = 0$.
- (ix) For each $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}' \in \tilde{\Omega}$, if $\mathbf{G}(\mathbf{x}_1)$ and $\mathbf{G}(\mathbf{x}_2)$ are linearly independent then $q_*(\mathbf{x}', \mathbf{x}_1)q_*(\mathbf{x}', \mathbf{x}_2) = 0$.
- (x) $q_* * \mathbf{G} \sim_{\text{post}} \mathbf{G}$.

Proof. The equivalences (i) \iff (ii), (iii) \iff (iv), and (ix) \iff (x) follow from Corollary 4 and the definition of the minimality. (i) \implies (v), (ii) \implies (iii), (v) \implies (x), and (vii) \implies (vi) are obvious. The equivalence (viii) \iff (ix) is immediate from the definition of q_* . The equivalence (v) \iff (vi) follows from Corollary 4 and that r given by (3) is in $K_{\mathbf{G}}^{\text{ind}}$ if and only if $r_\ell \in K(\mathbf{A}_\ell, \mathbf{G})$ for each $\ell \in \{1, \dots, n\}$.

(iii) \implies (ii). Assume (iii) and take an arbitrary element $p \in K_{\mathbf{G}}$. By the finite-dimensional Krein-Milman theorem, p is a convex combination of $\{p_1, \dots, p_N\}$, that is, there exists $(\mu_k)_{k=1}^N \in [0, 1]^N$ such that $\sum_{k=1}^N \mu_k = 1$ and $p = \sum_{k=1}^N \mu_k p_k$. Then

$$p(\mathbf{x}', \mathbf{x}) = \sum_{k=1}^N \mu_k p_k(\mathbf{x}', \mathbf{x}) \leq \sum_{k=1}^N p_k(\mathbf{x}', \mathbf{x}) = N p_*(\mathbf{x}', \mathbf{x}).$$

From this inequality and the assumption (iii), the condition (ii) follows.

(viii) \implies (vii) can be shown similarly.

(vi) \implies (ii). Assume that (ii) is not true. Then there exist $r \in K_{\mathbf{G}}$ and $\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2 \in \tilde{\Omega}$ such that $\mathbf{G}(\mathbf{x}_1)$ and $\mathbf{G}(\mathbf{x}_2)$ are linearly independent and $r(\mathbf{x}_0, \mathbf{x}_1)r(\mathbf{x}_0, \mathbf{x}_2) \neq 0$. We define $r_\ell \in K(\mathbf{A}_\ell, \mathbf{G})$ by

$$r_\ell(x'_\ell, \mathbf{x}) := \sum_{\mathbf{x}' \in \tilde{\Omega}, \pi_\ell(\mathbf{x}') = x'_\ell} r(\mathbf{x}', \mathbf{x}).$$

Then we have $r(\mathbf{x}', \mathbf{x}) \leq r_\ell(\pi_\ell(\mathbf{x}'), \mathbf{x})$ and hence

$$0 < (r(\mathbf{x}_0, \mathbf{x}_1)r(\mathbf{x}_0, \mathbf{x}_2))^n \leq \prod_{\ell=1}^n r_\ell(\pi_\ell(\mathbf{x}_0), \mathbf{x}_1)r_\ell(\pi_\ell(\mathbf{x}_0), \mathbf{x}_2).$$

Therefore we obtain $r_\ell(\pi_\ell(\mathbf{x}_0), \mathbf{x}_1)r_\ell(\pi_\ell(\mathbf{x}_0), \mathbf{x}_2) \neq 0$ for all $\ell \in \{1, \dots, n\}$, proving that (vi) is not true. \square

Theorem 2 gives the following two algorithms to determine whether a given joint observable \mathbf{G} is minimal or not. The first one is based on the condition (iii). According to Appendix C, the set extreme points $\{p_1, \dots, p_N\}$ of $K_{\mathbf{G}}$ can be explicitly calculated, so can be p_* . The other one is based on the condition (viii) or (ix), which can be explicitly checked by calculating the set of extreme points $\{q_{\ell,1}, \dots, q_{\ell,N_\ell}\}$ of the polytope $K(\mathbf{A}_\ell, \mathbf{F})$ for each ℓ .

At this point, we recall that an observable \mathbf{A} with an outcome set Ω is said to be *pairwise linearly independent* if any pair $(\mathbf{A}(x_1), \mathbf{A}(x_2))$, $x_1, x_2 \in \Omega$; $x_1 \neq x_2$, is linearly independent. Every observable is post-processing equivalent to a pairwise linearly independent observable unique up to the permutation of the outcome set [5, 11]. We refer to Appendix B for further details on this property. In the following we develop a method to determine the minimality of a joint observable which is pairwise linearly independent.

Firstly, let \mathbf{G} be a joint observable. For $u \in \mathbb{R}^{\Omega_\ell \times \tilde{\Omega}}$, let us consider the following system of homogeneous linear (in)equalities:

$$\sum_{\mathbf{x} \in \tilde{\Omega}} u(x'_\ell, \mathbf{x}) \mathbf{G}(\mathbf{x}) = 0 \quad (x'_\ell \in \Omega_\ell), \quad (4)$$

$$\sum_{x'_\ell \in \Omega_\ell} u(x'_\ell, \mathbf{x}) = 0 \quad (\mathbf{x} \in \tilde{\Omega}), \quad (5)$$

$$u(x'_\ell, \mathbf{x}) \begin{cases} \leq 0 & \text{if } x'_\ell = \pi_\ell(\mathbf{x}), \\ \geq 0 & \text{if } x'_\ell \neq \pi_\ell(\mathbf{x}). \end{cases} \quad (6)$$

We denote by $C_\ell(\mathbf{G})$ the pointed polyhedral cone on $\mathbb{R}^{\Omega_\ell \times \tilde{\Omega}}$ consisting of the solutions of (4), (5), and (6).

Proposition 2. *Suppose that \mathbf{G} is pairwise linearly independent. Then \mathbf{G} is a minimal joint observable if and only if $C_\ell(\mathbf{G}) = \{0\}$ for all $\ell \in \{1, \dots, n\}$.*

Proof. Assume that $C_\ell(\mathbf{G}) \neq \{0\}$ for some ℓ . Then we can take $u \in C_\ell(\mathbf{G})$, $x_\ell^0 \in \Omega_\ell$, and $\mathbf{x}_1 \in \tilde{\Omega}$ such that $u(x_\ell^0, \mathbf{x}_1) \neq 0$. If $x_\ell^0 = \pi_\ell(\mathbf{x}_1)$, then by (5) there exists another $x_\ell^1 \in \Omega_\ell$ satisfying $x_\ell^1 \neq x_\ell^0$ and $p(x_\ell^1, \mathbf{x}_1) \neq$

0. Thus, by also considering (6), we may assume $x_\ell^0 \neq \pi_\ell(\mathbf{x}_1)$ and $u(x_\ell^0, \mathbf{x}_1) > 0$. We define $\mathbf{x}_2 \in \tilde{\Omega}$ by

$$\pi_{\ell'}(\mathbf{x}_2) = \begin{cases} x_\ell^0 & \text{if } \ell' = \ell, \\ \pi_{\ell'}(\mathbf{x}_1) & \text{if } \ell' \neq \ell. \end{cases}$$

By multiplying a small positive constant if necessary, we may assume

$$|u(x'_\ell, \mathbf{x})| < 1 \quad \forall (x'_\ell, \mathbf{x}) \in \Omega_\ell \times \tilde{\Omega}. \quad (7)$$

We define $r_\ell \in \mathbb{R}^{\Omega_\ell \times \tilde{\Omega}}$ by

$$r_\ell(x'_\ell, \mathbf{x}) := \delta_{x'_\ell, \pi_\ell(\mathbf{x})} + u(x'_\ell, \mathbf{x}).$$

Then from (4), (5), (6), and (7), we have $r_\ell \in K(\mathbf{A}_\ell, \mathbf{G})$. Furthermore, we have $\mathbf{x}_1 \neq \mathbf{x}_2$ and

$$\begin{aligned} r_\ell(x_\ell^0, \mathbf{x}_1) &= u(x_\ell^0, \mathbf{x}_1) \neq 0, \\ r_\ell(x_\ell^0, \mathbf{x}_2) &= 1 + u(x_\ell^0, \mathbf{x}_2) \neq 0. \end{aligned}$$

We define $r \in \mathbf{Markov}(\tilde{\Omega}, \tilde{\Omega})$ by

$$r(\mathbf{x}', \mathbf{x}) := r_\ell(\pi_\ell(\mathbf{x}'), \mathbf{x}) \prod_{\ell' \neq \ell} \delta_{\pi_{\ell'}(\mathbf{x}'), \pi_{\ell'}(\mathbf{x})}.$$

Then $r * \mathbf{G} \in \mathcal{J}(\{\mathbf{A}_{\ell'}\}_{\ell'=1}^n)$ and

$$\begin{aligned} r(\mathbf{x}_2, \mathbf{x}_1) &= r_\ell(x_\ell^0, \mathbf{x}_1) \neq 0, \\ r(\mathbf{x}_2, \mathbf{x}_2) &= r_\ell(x_\ell^0, \mathbf{x}_2) \neq 0. \end{aligned}$$

This implies that the condition (ii) of Theorem 2 is not true. Therefore \mathbf{G} is not minimal.

Assume that $C_\ell(\mathbf{G}) = \{0\}$ for all $\ell \in \{1, \dots, n\}$. Let $\mathbf{G}' \in \mathcal{J}(\{\mathbf{A}_{\ell'}\}_{\ell'=1}^n)$ be a joint observable satisfying $\mathbf{G}' \preceq_{\text{post}} \mathbf{G}$. Then we can take $r \in \mathbf{Markov}(\tilde{\Omega}, \tilde{\Omega})$ such that $\mathbf{G}' = r * \mathbf{G}$. For each ℓ we define $r_\ell \in K(\mathbf{A}_\ell, \mathbf{G})$ by

$$r_\ell(x'_\ell, \mathbf{x}) := \sum_{\mathbf{x}', \pi_\ell(\mathbf{x}') = x'_\ell} r(\mathbf{x}', \mathbf{x}),$$

and $u_\ell \in \mathbb{R}^{\Omega_\ell \times \tilde{\Omega}}$ by

$$u_\ell(x'_\ell, \mathbf{x}) := r_\ell(x'_\ell, \mathbf{x}) - \delta_{x'_\ell, \pi_\ell(\mathbf{x})}.$$

From $r_\ell(x'_\ell, \mathbf{x}) \in [0, 1]$ we can easily check that $u_\ell \in C_\ell(\mathbf{G})$. Therefore by assumption we have $u_\ell = 0$ and hence $r_\ell(x'_\ell, \mathbf{x}) = \delta_{x'_\ell, \pi_\ell(\mathbf{x})}$. If $\mathbf{x}' \neq \mathbf{x}$, there exists ℓ with $\pi_\ell(\mathbf{x}') \neq \pi_\ell(\mathbf{x})$. Thus

$$r(\mathbf{x}', \mathbf{x}) \leq r_\ell(\pi_\ell(\mathbf{x}'), \mathbf{x}) = \delta_{\pi_\ell(\mathbf{x}'), \pi_\ell(\mathbf{x})} = 0.$$

This implies $r(\mathbf{x}', \mathbf{x}) = \delta_{\mathbf{x}, \mathbf{x}'}$ and hence $\mathbf{G}' = r * \mathbf{G} = \mathbf{G}$. Thus \mathbf{G} is minimal. \square

For each ℓ , $C_\ell(\mathbf{G})$ is a polyhedral cone and is the conical hull of a finite set $\{u_{\ell 1}, \dots, u_{\ell, M_\ell}\}$, which can be calculated according to Appendix C. Thus Proposition 2 provides a method to determine the minimality of \mathbf{G} when \mathbf{G} is pairwise linearly independent.

Proposition 3. *Let $\tilde{\Omega}_1 := \{\mathbf{x} \in \tilde{\Omega} \mid \mathbf{G}(\mathbf{x}) \neq 0\}$. Suppose that $(\mathbf{G}(\mathbf{x}))_{\mathbf{x} \in \tilde{\Omega}_1}$ is linearly independent. Then \mathbf{G} is a minimal joint observable.*

Proof. Let $\mathbf{G}' \in \mathcal{J}(\{\mathbf{A}_\ell\}_{\ell=1}^n)$ be a joint observable satisfying $\mathbf{G}' \preceq_{\text{post}} \mathbf{G}$ and let $p \in \mathbf{Markov}(\tilde{\Omega}, \tilde{\Omega})$ be a Markov kernel such that $\mathbf{G}' = p * \mathbf{G}$. From $\mathbf{G}, \mathbf{G}' \in \mathcal{J}(\{\mathbf{A}_\ell\}_{\ell=1}^n)$ we have

$$\begin{aligned} \sum_{\mathbf{x} \in \tilde{\Omega}_1, \pi_\ell(\mathbf{x}) = x_\ell} \mathbf{G}(\mathbf{x}) &= \mathbf{A}_\ell(x_\ell) = \sum_{\mathbf{x}' \in \tilde{\Omega}, \pi_\ell(\mathbf{x}') = x_\ell} \mathbf{G}'(\mathbf{x}') \\ &= \sum_{\mathbf{x} \in \tilde{\Omega}_1} \left(\sum_{\mathbf{x}' \in \tilde{\Omega}, \pi_\ell(\mathbf{x}') = x_\ell} p(\mathbf{x}', \mathbf{x}) \right) \mathbf{G}(\mathbf{x}). \end{aligned}$$

Thus the linear independence of $(\mathbf{G}(\mathbf{x}))_{\mathbf{x} \in \tilde{\Omega}_1}$ implies

$$\sum_{\mathbf{x}', \pi_\ell(\mathbf{x}') = x_\ell} p(\mathbf{x}', \mathbf{x}) = \delta_{\pi_\ell(\mathbf{x}), x_\ell}$$

for all $\mathbf{x} \in \tilde{\Omega}_1$. By the same discussion as in Proposition 2, we obtain $p(\mathbf{x}', \mathbf{x}) = \delta_{\mathbf{x}, \mathbf{x}'}$ ($\forall \mathbf{x} \in \tilde{\Omega}_1, \forall \mathbf{x}' \in \tilde{\Omega}$) and hence $\mathbf{G} = \mathbf{G}'$. Therefore, \mathbf{G} is minimal. \square

Corollary 2. *Suppose that \mathbf{G} is linearly independent. Then \mathbf{G} is both minimal and maximal.*

Proof. The minimality of \mathbf{G} is immediate from Proposition 3.

Assume $\mathbf{G} \preceq_{\text{post}} \mathbf{G}'$ for some $\mathbf{G}' \in \mathcal{J}(\{\mathbf{A}_\ell\}_{\ell=1}^n)$. Since each element of \mathbf{G} is a linear combination of \mathbf{G}' , a dimensional argument yields that \mathbf{G}' is also linearly independent. (Here we used the finiteness of the outcome set $\tilde{\Omega}$). Hence from the proof of Proposition 3 we obtain $\mathbf{G} = \mathbf{G}'$. Therefore, \mathbf{G} is maximal. \square

5. DICHOTOMIC QUBIT OBSERVABLES

5.1. Compatibility of dichotomic qubit observables. We denote $\boldsymbol{\sigma} = (\sigma_1, \sigma_2, \sigma_3)$. A dichotomic qubit observable is of the form

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

where $\mathbf{a} \in \mathbb{R}^3$, $\alpha \in [0, 2]$ and $\|\mathbf{a}\| \leq \min(\alpha, 2 - \alpha)$.

All joint observables of a compatible pair $(\mathbf{E}^{\alpha,\mathbf{a}}, \mathbf{E}^{\beta,\mathbf{b}})$ are parametrized by two parameters $\gamma \in \mathbb{R}$ and $\mathbf{g} \in \mathbb{R}^3$ in the following way. The joint observable $\mathbf{G}^{\gamma,\mathbf{g}}$ is defined as

$$\begin{aligned} \mathbf{G}^{\gamma,\mathbf{g}}(+, +) &= \frac{1}{2}(\gamma\mathbb{1} + \mathbf{g} \cdot \boldsymbol{\sigma}), \\ \mathbf{G}^{\gamma,\mathbf{g}}(+, -) &= \mathbf{E}^{1,\mathbf{a}}(+)-\mathbf{G}^{\gamma,\mathbf{g}}(+, +) \\ &= \frac{1}{2}[(\alpha-\gamma)\mathbb{1} + (\mathbf{a}-\mathbf{g}) \cdot \boldsymbol{\sigma}], \\ \mathbf{G}^{\gamma,\mathbf{g}}(-, +) &= \mathbf{E}^{1,\mathbf{b}}(+)-\mathbf{G}^{\gamma,\mathbf{g}}(+, +) \\ &= \frac{1}{2}[(\beta-\gamma)\mathbb{1} + (\mathbf{b}-\mathbf{g}) \cdot \boldsymbol{\sigma}] \\ \mathbf{G}^{\gamma,\mathbf{g}}(-, -) &= \mathbb{1} + \mathbf{G}^{\gamma,\mathbf{g}}(+, +) - \mathbf{E}^{1,\mathbf{a}}(+)-\mathbf{E}^{1,\mathbf{b}}(+), \\ &= \frac{1}{2}[(2+\gamma-\alpha-\beta)\mathbb{1} + (\mathbf{g}-\mathbf{a}-\mathbf{b}) \cdot \boldsymbol{\sigma}]. \end{aligned}$$

For $\mathbf{G}^{\gamma,\mathbf{g}}$ to be a valid joint observable, these four operators must be positive. This means that the parameters γ, \mathbf{g} have to satisfy the following inequalities [12]:

$$\|\mathbf{g}\| \leq \gamma \tag{8}$$

$$\|\mathbf{a}-\mathbf{g}\| \leq \alpha-\gamma \tag{9}$$

$$\|\mathbf{b}-\mathbf{g}\| \leq \beta-\gamma \tag{10}$$

$$\|\mathbf{a}+\mathbf{b}-\mathbf{g}\| \leq 2+\gamma-\alpha-\beta \tag{11}$$

From the previous inequalities one can solve the compatibility condition for $\mathbf{E}^{\alpha,\mathbf{a}}$ and $\mathbf{E}^{\beta,\mathbf{b}}$, and three equivalent formulations are presented in [12, 13, 14]. For our purposes, we do not need the compatibility condition; we simply assume that $\mathbf{E}^{\alpha,\mathbf{a}}$ and $\mathbf{E}^{\beta,\mathbf{b}}$ are compatible and in the following we analyze the condition for the minimality of joint observable.

5.2. Minimal joint observables. Now we give a complete characterization of the minimality of $\mathbf{G}^{\gamma,\mathbf{g}}$ when the marginal observables $\mathbf{E}^{\alpha,\mathbf{a}}$ and $\mathbf{E}^{\beta,\mathbf{b}}$ are non-commutative, which holds if and only if \mathbf{a} and \mathbf{b} are linearly independent. We put $\mathbf{A}_1 := \mathbf{E}^{\alpha,\mathbf{a}}$ and $\mathbf{A}_2 := \mathbf{E}^{\beta,\mathbf{b}}$ and adopt the same notation as in Section 4.

5.2.1. Trivial compatibility. Two dichotomic observables \mathbf{A}_1 and \mathbf{A}_2 are trivially compatible if one of the orderings $\mathbf{A}_1(+) \leq \mathbf{A}_2(+)$, $\mathbf{A}_1(+) \geq \mathbf{A}_2(-)$, $\mathbf{A}_1(-) \leq \mathbf{A}_2(-)$ or $\mathbf{A}_1(-) \geq \mathbf{A}_2(+)$ holds [15]. Two compatible dichotomic observables satisfy this kind of trivial condition exactly when they have a joint observable with one element being zero.

Suppose that $\mathbf{G}^{\gamma;\mathbf{g}}(++) = 0$, which is equivalent to $\mathbf{A}_1(+) \leq \mathbf{A}_2(-)$. In this case, the other elements of $\mathbf{G}^{\gamma;\mathbf{g}}$ are determined to be

$$\begin{aligned}\mathbf{G}^{\gamma;\mathbf{g}}(+ -) &= \mathbf{A}_1(+), \\ \mathbf{G}^{\gamma;\mathbf{g}}(- +) &= \mathbf{A}_2(+), \\ \mathbf{G}^{\gamma;\mathbf{g}}(- -) &= \mathbb{1} - \mathbf{A}_1(+) - \mathbf{A}_2(+).\end{aligned}$$

Analogous equations follow in other cases when $\mathbf{G}^{\gamma;\mathbf{g}}(+ -) = 0$, $\mathbf{G}^{\gamma;\mathbf{g}}(- +) = 0$ or $\mathbf{G}^{\gamma;\mathbf{g}}(- -) = 0$.

We conclude that in the case of trivial compatibility, a joint measurement has three nonzero elements and it is unique.

5.2.2. Linearly independent vectors. Secondly, we consider the case when the vectors $\mathbf{a}, \mathbf{b}, \mathbf{g}$ are linearly independent. In this case, we can easily check that $\mathbf{G}^{\gamma;\mathbf{g}}$ is linearly independent. Hence Corollary 2 implies that $\mathbf{G}^{\gamma;\mathbf{g}}$ is both maximal and minimal.

5.2.3. Linearly dependent vectors. Thirdly, we consider the case when \mathbf{g} can be written as a linear combination $\mathbf{g} = c_1\mathbf{a} + c_2\mathbf{b}$, $c_1, c_2 \in \mathbb{R}$. In order to calculate $K(\mathbf{A}_\ell, \mathbf{G}^{\gamma;\mathbf{g}})$ and $C_\ell(\mathbf{G}^{\gamma;\mathbf{g}})$, consider the following homogeneous linear (in)equalities for $u \in \mathbb{R}^{\tilde{\Omega}}$:

$$\begin{pmatrix} \gamma & \alpha - \gamma & \beta - \gamma & 2 + \gamma - \alpha - \beta \\ c_1 & 1 - c_1 & -c_1 & c_1 - 1 \\ c_2 & -c_2 & 1 - c_2 & c_2 - 1 \end{pmatrix} \begin{pmatrix} u(++) \\ u(+ -) \\ u(- +) \\ u(- -) \end{pmatrix} = 0, \quad (12)$$

$$u(x, x') \begin{cases} \leq 0 & \text{if } x = + \\ \geq 0 & \text{if } x = -, \end{cases} \quad (13)$$

$$u(x, x') \begin{cases} \leq 0 & \text{if } x' = + \\ \geq 0 & \text{if } x' = -. \end{cases} \quad (14)$$

Let $C_{1,+}$ be the polyhedral cone on $\mathbb{R}^{\tilde{\Omega}}$ defined by (12) and (13), and $C_{2,+}$ be the one defined by (12) and (14). Then it can be checked that $K(\mathbf{A}_\ell, \mathbf{G}^{\gamma;\mathbf{g}})$ and $C_\ell(\mathbf{G}^{\gamma;\mathbf{g}})$ can be given as follows.

- For $p \in \mathbb{R}^{\Omega_\ell \times \tilde{\Omega}}$, $p \in K(\mathbf{A}_\ell, \mathbf{G}^{\gamma;\mathbf{g}})$ if and only if there exists $u_+ \in C_{1,+}$ such that $p(+, \mathbf{x}) = \delta_{+, \pi_\ell(\mathbf{x})} + u_+(\mathbf{x})$ and $p(-, \mathbf{x}) = \delta_{-, \pi_\ell(\mathbf{x})} - u_+(\mathbf{x}) \geq 0$ ($\mathbf{x} \in \tilde{\Omega}$).
- For $v \in \mathbb{R}^{\Omega_\ell \times \tilde{\Omega}}$, $v \in C_\ell(\mathbf{G}^{\gamma;\mathbf{g}})$ if and only if there exists $u_+ \in C_{2,+}$ such that $v(+, \mathbf{x}) = u_+(\mathbf{x})$ and $v(-, \mathbf{x}) = -u_+(\mathbf{x})$ ($\mathbf{x} \in \tilde{\Omega}$).

The general solution of (12) is $u = tw$ ($t \in \mathbb{R}$), where $w \in \mathbb{R}^{\tilde{\Omega}}$ is given by

$$\begin{pmatrix} w(++), \\ w(+-), \\ w(-+), \\ w(--) \end{pmatrix} := \begin{pmatrix} (2 - \alpha)c_1 + (2 - \beta)c_2 + \gamma - 2 \\ (2 - \alpha)c_1 - \beta c_2 + \gamma \\ -\alpha c_1 + (2 - \beta)c_2 + \gamma \\ -\alpha c_1 - \beta c_2 + \gamma \end{pmatrix}. \quad (15)$$

If $\mathbf{G}^{\gamma, \mathbf{g}}$ is pairwise linearly independent (see Appendix B), then by Proposition 2, $\mathbf{G}^{\gamma, \mathbf{g}}$ is a minimal joint observable if and only if $C_1(\mathbf{G}^{\gamma, \mathbf{g}}) = C_2(\mathbf{G}^{\gamma, \mathbf{g}}) = \{0\}$, or equivalently, $C_{1,+} = C_{2,+} = \{0\}$. By noting $w \neq 0$, $C_{1,+} \neq \{0\}$ if and only if

$$\begin{aligned} & [w(++), \leq 0 \wedge w(+-, \leq 0 \wedge w(-+, \geq 0 \wedge w(--), \geq 0] \\ & \vee [w(++), \geq 0 \wedge w(+-, \geq 0 \wedge w(-+, \leq 0 \wedge w(--), \leq 0] \end{aligned}$$

Hence $C_{1,+} = \{0\}$ if and only if

$$\begin{aligned} & [w(++), > 0 \vee w(+-, > 0 \vee w(-+, < 0 \vee w(--), < 0] \\ & \wedge [w(++), < 0 \vee w(+-, < 0 \vee w(-+, > 0 \vee w(--), > 0]. \end{aligned}$$

By using the equivalences

$$\begin{aligned} (x > 0 \wedge y > 0) \vee (x < 0 \wedge y < 0) & \iff xy > 0, \\ (x > 0 \wedge y < 0) \vee (x < 0 \wedge y > 0) & \iff xy < 0 \end{aligned}$$

valid for $x, y \in \mathbb{R}$, we obtain

$$\begin{aligned} & C_{1,+} = \{0\} \\ \iff & (w(++), w(+-, < 0) \vee (w(++), w(-+, > 0) \vee (w(++), w(--), > 0) \\ & \vee (w(+-, w(-+, > 0) \vee (w(+-, w(--), > 0) \vee (w(-+, w(--), < 0) \end{aligned}$$

Similarly we obtain

$$\begin{aligned} & C_{2,+} = \{0\} \\ \iff & (w(++), w(+-, > 0) \vee (w(++), w(-+, < 0) \vee (w(++), w(--), > 0) \\ & \vee (w(+-, w(-+, > 0) \vee (w(+-, w(--), < 0) \vee (w(-+, w(--), > 0). \end{aligned}$$

Hence

$$\begin{aligned}
& C_{1,+} = C_{2,+} = \{0\} \\
\iff & (w(++)w(--) > 0) \vee (w(+-)w(-+) > 0) \\
& \vee \{[(w(++)w(+-) < 0) \vee (w(++)w(-+) > 0) \vee (w(+-)w(--) > 0) \\
& \vee (w(-+)w(--) < 0)] \wedge [(w(++)w(+-) > 0) \vee (w(++)w(-+) < 0) \\
& \vee (w(+-)w(--) < 0) \vee (w(-+)w(--) > 0)]\} \\
\iff & (w(++)w(--) > 0) \vee (w(+-)w(-+) > 0) \\
& \vee (w(++)^2 w(+-)w(-+) > 0) \vee (w(++)w(+-)^2 w(--) > 0) \\
& \vee (w(++)w(+-)w(-+)w(--) < 0) \vee (w(++)w(-+)^2 w(--) > 0) \\
& \vee (w(+-)w(-+)w(--)^2 > 0) \\
\iff & (w(++)w(--) > 0) \vee (w(+-)w(-+) > 0) \vee (w(++)w(+-)w(-+)w(--) < 0) \\
\iff & (w(++)w(--) > 0) \vee (w(+-)w(-+) > 0) \tag{16}
\end{aligned}$$

Now we consider the case when $\mathbf{G}^{\gamma, \mathbf{g}}$ is not pairwise linearly independent. The linear dependence conditions for the pairs of the elements of $\mathbf{G}^{\gamma, \mathbf{g}}$ are given as follows.

(i) $\mathbf{G}^{\gamma, \mathbf{g}}(++)$ and $\mathbf{G}^{\gamma, \mathbf{g}}(+-)$ are linearly dependent if and only if

$$\mathbf{g} = \frac{\gamma}{\alpha} \mathbf{a}. \tag{17}$$

(ii) $\mathbf{G}^{\gamma, \mathbf{g}}(++)$ and $\mathbf{G}^{\gamma, \mathbf{g}}(-+)$ are linearly dependent if and only if

$$\mathbf{g} = \frac{\gamma}{\beta} \mathbf{b}. \tag{18}$$

(iii) $\mathbf{G}^{\gamma, \mathbf{g}}(++)$ and $\mathbf{G}^{\gamma, \mathbf{g}}(--)$ are linearly dependent if and only if

$$\alpha + \beta \neq 2 \wedge \mathbf{g} = \frac{\gamma}{\alpha + \beta - 2} (\mathbf{a} + \mathbf{b}). \tag{19}$$

(iv) $\mathbf{G}^{\gamma, \mathbf{g}}(+-)$ and $\mathbf{G}^{\gamma, \mathbf{g}}(-+)$ are linearly dependent if and only if

$$\alpha \neq \beta \wedge \mathbf{g} = \frac{\gamma - \beta}{\alpha - \beta} \mathbf{a} + \frac{\gamma - \alpha}{\beta - \alpha} \mathbf{b}. \tag{20}$$

Here, the possibility of $\alpha = \beta$ will be excluded since this implies $\mathbf{g} = \mathbf{a} = \mathbf{b}$, contradicting the linear independence of \mathbf{a} and \mathbf{b} .

(v) $\mathbf{G}^{\gamma, \mathbf{g}}(+-)$ and $\mathbf{G}^{\gamma, \mathbf{g}}(--)$ are linearly dependent if and only if

$$\beta \neq 2 \wedge \mathbf{g} = \mathbf{a} + \frac{\alpha - \gamma}{2 - \beta} \mathbf{b}. \tag{21}$$

Here, the possibility of $\beta = 2$ will be excluded since this implies $\mathbf{a} = \mathbf{g} = \mathbf{a} + \mathbf{b}$, contradicting $\mathbf{b} \neq 0$.

(vi) $\mathbf{G}^{\gamma, \mathbf{g}}(-+)$ and $\mathbf{G}^{\gamma, \mathbf{g}}(--)$ are linearly dependent if and only if

$$\alpha \neq 2 \wedge \mathbf{g} = \frac{\beta - \gamma}{2 - \alpha} \mathbf{a} + \mathbf{b}. \quad (22)$$

Here, the possibility of $\alpha = 2$ will be excluded since this implies $\mathbf{b} = \mathbf{g} = \mathbf{a} + \mathbf{b}$, contradicting $\mathbf{a} \neq 0$.

As a conclusion, $\mathbf{G}^{\gamma, \mathbf{g}}$ is pairwise linearly independent if and only if neither of the above six conditions holds. We remark that each of the above conditions corresponds to the intersection point of two lines $w(\mathbf{x}) = 0$ and $w(\mathbf{x}') = 0$ on c_1 - c_2 plane. For example, the condition (17) holds if and only if $w(-+) = w(--)= 0$.

Now assume that $\mathbf{G}^{\gamma, \mathbf{g}}(++)$ and $\mathbf{G}^{\gamma, \mathbf{g}}(+)$ are linearly dependent. Then $\mathbf{g} = \frac{\gamma}{\alpha} \mathbf{a}$, and $c_1 = \frac{\gamma}{\alpha}$, $c_2 = 0$. Hence

$$w(++)= -2\frac{\alpha - \gamma}{\alpha}, w(+)= 2\frac{\gamma}{\alpha}, w(-+)= w(--)= 0.$$

Thus each $r_\ell \in K(\mathbf{A}_\ell, \mathbf{G}^{\gamma, \mathbf{g}})$ ($\ell = 1, 2$) can be written as

$$\begin{pmatrix} r_1(+, ++) \\ r_1(+, +) \\ r_1(+, -) \\ r_1(+, --) \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} r_1(-, ++) \\ r_1(-, +) \\ r_1(-, -) \\ r_1(-, --) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix},$$

$$\begin{pmatrix} r_2(+, ++) \\ r_2(+, +) \\ r_2(+, -) \\ r_2(+, --) \end{pmatrix} = \begin{pmatrix} 1 - 2s\frac{\alpha - \gamma}{\alpha} \\ 2s\frac{\gamma}{\alpha} \\ 1 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} r_2(-, ++) \\ r_2(-, +) \\ r_2(-, -) \\ r_2(-, --) \end{pmatrix} = \begin{pmatrix} 2s\frac{\alpha - \gamma}{\alpha} \\ 1 - 2s\frac{\gamma}{\alpha} \\ 0 \\ 1 \end{pmatrix}$$

for some $s \geq 0$. We can easily check the condition (vi) of Theorem 2, which implies the minimality of $\mathbf{G}^{\gamma, \mathbf{g}}$.

We can similarly check the minimality of $\mathbf{G}^{\gamma, \mathbf{g}}$ when either $(\mathbf{G}^{\gamma, \mathbf{g}}(++), \mathbf{G}^{\gamma, \mathbf{g}}(-+))$, $(\mathbf{G}^{\gamma, \mathbf{g}}(+), \mathbf{G}^{\gamma, \mathbf{g}}(--))$, or $(\mathbf{G}^{\gamma, \mathbf{g}}(-+), \mathbf{G}^{\gamma, \mathbf{g}}(--))$ is a linearly dependent pair.

Assume that $\mathbf{G}^{\gamma, \mathbf{g}}(++)$ and $\mathbf{G}^{\gamma, \mathbf{g}}(--)$ are linearly dependent. Then $\alpha + \beta \neq 2$ and $\mathbf{g} = \frac{\gamma}{\alpha + \beta - 2}(\mathbf{a} + \mathbf{b})$. Thus

$$w(++)= -2\frac{2 + \gamma - \alpha - \beta}{2 - \alpha - \beta}, w(-+)= w(+)= 0, w(--)= \frac{2\gamma}{2 - \alpha - \beta}.$$

Hence each $r_\ell \in K(\mathbf{A}_\ell, \mathbf{G}^{\gamma, \mathbf{g}})$ ($\ell = 1, 2$) can be written as

$$\begin{pmatrix} r_1(+, ++) \\ r_1(+, +-) \\ r_1(+, -+) \\ r_1(+, --) \end{pmatrix} = \begin{pmatrix} 1 - t(2 + \gamma - \alpha - \beta) \\ 1 \\ 0 \\ t\gamma \end{pmatrix}, \quad \begin{pmatrix} r_1(-, ++) \\ r_1(-, +-) \\ r_1(-, -+) \\ r_1(-, --) \end{pmatrix} = \begin{pmatrix} t(2 + \gamma - \alpha - \beta) \\ 0 \\ 1 \\ 1 - t\gamma \end{pmatrix},$$

$$\begin{pmatrix} r_2(+, ++) \\ r_2(+, +-) \\ r_2(+, -+) \\ r_2(+, --) \end{pmatrix} = \begin{pmatrix} 1 - s(2 + \gamma - \alpha - \beta) \\ 0 \\ 1 \\ s\gamma \end{pmatrix}, \quad \begin{pmatrix} r_2(-, ++) \\ r_2(-, +-) \\ r_2(-, -+) \\ r_2(-, --) \end{pmatrix} = \begin{pmatrix} s(2 + \gamma - \alpha - \beta) \\ 1 \\ 0 \\ 1 - s\gamma \end{pmatrix}$$

for some $t, s \geq 0$. By taking sufficiently small $t, s > 0$, we have

$$r_1(+, ++)r_1(+, +-) \neq 0 \neq r_2(-, ++)r_2(-, +-).$$

Since $(\mathbf{G}^{\gamma, \mathbf{g}}(+), \mathbf{G}^{\gamma, \mathbf{g}}(+))$ is a linearly independent pair in this case, Theorem 2 implies that $\mathbf{G}^{\gamma, \mathbf{g}}$ is not minimal.

Assume that $\mathbf{G}^{\gamma, \mathbf{g}}(+)$ and $\mathbf{G}^{\gamma, \mathbf{g}}(-)$ are linearly dependent. Then $\alpha \neq \beta$ and $\mathbf{g} = \frac{\gamma - \beta}{\alpha - \beta} \mathbf{a} + \frac{\gamma - \alpha}{\beta - \alpha} \mathbf{b}$. Thus

$$w(++) = w(--) = 0, \quad w(+ -) = 2 \frac{\beta - \gamma}{\beta - \alpha}, \quad w(- +) = 2 \frac{\alpha - \gamma}{\alpha - \beta}.$$

Hence each $r_\ell \in K(\mathbf{A}_\ell, \mathbf{G}^{\gamma, \mathbf{g}})$ ($\ell = 1, 2$) can be written as

$$\begin{pmatrix} r_1(+, ++) \\ r_1(+, +-) \\ r_1(+, -+) \\ r_1(+, --) \end{pmatrix} = \begin{pmatrix} 1 \\ 1 - t(\beta - \gamma) \\ t(\alpha - \gamma) \\ 0 \end{pmatrix}, \quad \begin{pmatrix} r_1(-, ++) \\ r_1(-, +-) \\ r_1(-, -+) \\ r_1(-, --) \end{pmatrix} = \begin{pmatrix} 0 \\ t(\beta - \gamma) \\ 1 - t(\alpha - \gamma) \\ 1 \end{pmatrix},$$

$$\begin{pmatrix} r_2(+, ++) \\ r_2(+, +-) \\ r_2(+, -+) \\ r_2(+, --) \end{pmatrix} = \begin{pmatrix} 1 \\ s(\beta - \gamma) \\ 1 - s(\alpha - \gamma) \\ 0 \end{pmatrix}, \quad \begin{pmatrix} r_2(-, ++) \\ r_2(-, +-) \\ r_2(-, -+) \\ r_2(-, --) \end{pmatrix} = \begin{pmatrix} 0 \\ 1 - s(\beta - \gamma) \\ s(\alpha - \gamma) \\ 1 \end{pmatrix}$$

for some $t, s \geq 0$. By taking sufficiently small $t, s > 0$, we have

$$r_1(+, ++)r_1(+, +-) \neq 0 \neq r_2(+, ++)r_2(+, +-).$$

Since $\mathbf{G}^{\gamma, \mathbf{g}}(+)$ and $\mathbf{G}^{\gamma, \mathbf{g}}(-)$ is linearly independent in this case, Theorem 2 implies that $\mathbf{G}^{\gamma, \mathbf{g}}$ is not minimal.

We remark that the above two non-minimal cases (19) and (20) are inconsistent with the condition (16).

To summarize, we obtain the following proposition which completely characterizes the minimality $\mathbf{G}^{\gamma, \mathbf{g}}$.

Theorem 3. *Let $\mathbf{E}^{\alpha, \mathbf{a}}$ and $\mathbf{E}^{\beta, \mathbf{b}}$ be compatible qubit observables. Suppose that \mathbf{a} and \mathbf{b} are linearly independent. Then their joint observable*

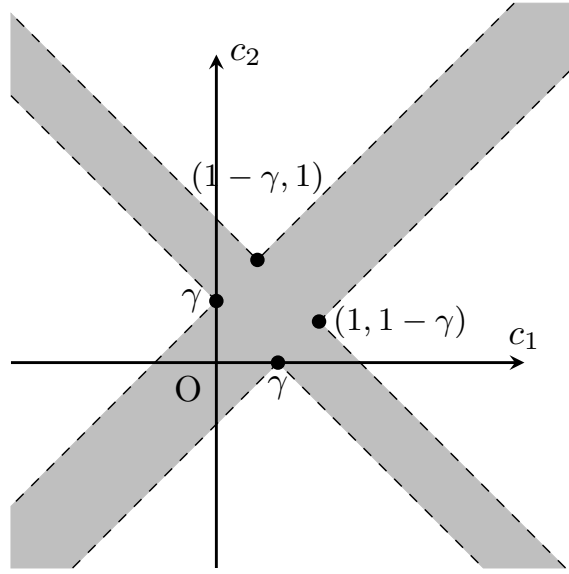


FIGURE 1. The region of (c_1, c_2) such that $\mathbf{G}^{\gamma, \mathbf{g}}$ is a minimal joint observable when $\mathbf{g} = c_1 \mathbf{a} + c_2 \mathbf{b}$. Note that \mathbf{g} should also satisfy the positivity conditions (8)–(11) of $\mathbf{G}^{\gamma, \mathbf{g}}$, whose shape in the c_1 - c_2 plane strongly depends on the values of the inner products $\mathbf{a} \cdot \mathbf{a}$, $\mathbf{b} \cdot \mathbf{b}$, and $\mathbf{a} \cdot \mathbf{b}$.

$\mathbf{G}^{\gamma, \mathbf{g}}$ is a minimal joint observable if and only if either of the following conditions holds.

1. \mathbf{g} , \mathbf{a} , and \mathbf{b} are linearly independent.
2. One (and only one) element of $\mathbf{G}^{\gamma, \mathbf{g}}$ is zero.
3. $\mathbf{G}^{\gamma, \mathbf{g}}$ satisfies either of the pairwise linear dependence conditions (17), (18), (21), and (22).
4. \mathbf{g} can be written as $\mathbf{g} = c_1 \mathbf{a} + c_2 \mathbf{b}$ ($c_1, c_2 \in \mathbb{R}$) and satisfies (16).

5.3. Unbiased dichotomic qubit observables. Finally, we consider the simple special case $\alpha = \beta = 1$. These kind of observables are called *unbiased*. As shown in [16], $\mathbf{E}^{1, \mathbf{a}}$ and $\mathbf{E}^{1, \mathbf{b}}$ are compatible if and only if

$$\|\mathbf{a} - \mathbf{b}\| + \|\mathbf{a} + \mathbf{b}\| \leq 2. \quad (23)$$

If \mathbf{a} and \mathbf{b} are nonzero different vectors, then we easily see that two unbiased observables $\mathbf{E}^{1, \mathbf{a}}$ and $\mathbf{E}^{1, \mathbf{b}}$ cannot be trivially compatible. In particular, if \mathbf{a} and \mathbf{b} are linearly independent, then all elements of $\mathbf{G}^{\gamma, \mathbf{g}}$ must be nonzero.

The notation (15) now takes the form

$$\begin{pmatrix} w(++) \\ w(+-) \\ w(-+) \\ w(--) \end{pmatrix} = \begin{pmatrix} c_1 + c_2 + \gamma - 2 \\ c_1 - c_2 + \gamma \\ -c_1 + c_2 + \gamma \\ -c_1 - c_2 + \gamma \end{pmatrix}$$

and the condition (16) reduces to

$$(\gamma < c_1 + c_2 < 2 - \gamma) \vee (-\gamma < c_1 - c_2 < \gamma). \quad (24)$$

We also note that $0 < \gamma < 1$ in this case. In fact, if $\gamma = 0$, we have $\mathbf{G}^{\gamma, \mathbf{g}}(++) = \mathbf{G}^{\gamma, \mathbf{g}}(--) = 0$ which contradicts the linear independence of \mathbf{a} and \mathbf{b} . If $\gamma = 1$, we have $\mathbf{G}^{\gamma, \mathbf{g}}(+-) = \mathbf{G}^{\gamma, \mathbf{g}}(-+) = 0$, which again contradicts the linear independence of \mathbf{a} and \mathbf{b} . Thus we obtain the following corollary.

Corollary 3. *Let $\mathbf{E}^{1, \mathbf{a}}$ and $\mathbf{E}^{1, \mathbf{b}}$ be compatible qubit observables and suppose that \mathbf{a} and \mathbf{b} are linearly independent. Then their joint observable $\mathbf{G}^{\gamma, \mathbf{g}}$ is a minimal joint observable if and only if one of the following conditions holds.*

1. \mathbf{g} , \mathbf{a} , and \mathbf{b} are linearly independent.
2. $\mathbf{G}^{\gamma, \mathbf{g}}$ satisfies either of the pairwise linear dependence conditions (17), (18), (21), and (22).
3. \mathbf{g} can be written as $\mathbf{g} = c_1 \mathbf{a} + c_2 \mathbf{b}$ ($c_1, c_2 \in \mathbb{R}$) and satisfies (24).

For a fixed $\gamma \in (0, 1)$, the region for $\mathbf{g} = c_1 \mathbf{a} + c_2 \mathbf{b}$ corresponding to the conditions 2 and 3 of Corollary 3 is depicted in Fig. 1.

ACKNOWLEDGEMENTS

T.H. acknowledges financial support from the Academy of Finland (Project No. 287750) and from the Academy of Finland Centre of Excellence program (Project No. 312058). Y.K. acknowledges financial support by the National Natural Science Foundation of China (Grants No. 11374375 and No. 11574405).

REFERENCES

- [1] T. Heinosaari, T. Miyadera, and M. Ziman. An invitation to quantum incompatibility. *J. Phys. A: Math. Theor.*, 49:123001, 2016.
- [2] T. Heinosaari, D. Reitzner, and P. Stano. Notes on joint measurability of quantum observables. *Found. Phys.*, 38:1133–1147, 2008.
- [3] E. Haapasalo, T. Heinosaari, and J.-P. Pellonpää. When do pieces determine the whole? Extreme marginals of a completely positive map. *Rev. Math. Phys.*, 26:1450002, 2014.
- [4] L. Guerini and M.T. Cunha. Uniqueness of the joint measurement and the structure of the set of compatible quantum measurements. *J. Math. Phys.*, 59:042106, 2018.
- [5] H. Martens and W.M. de Muynck. Nonideal quantum measurements. *Found. Phys.*, 20:255–281, 1990.
- [6] F. Buscemi, G.M. D’Ariano, M. Keyl, P. Perinotti, and R.F. Werner. Clean positive operator valued measures. *J. Math. Phys.*, 46:082109, 2005.
- [7] T. Heinonen. Optimal measurements in quantum mechanics. *Phys. Lett. A*, 346:77–86, 2005.
- [8] T. Heinosaari and T. Miyadera. Qualitative noise-disturbance relation for quantum measurements. *Phys. Rev. A*, 88:042117, 2013.
- [9] S.T. Ali, C. Carmeli, T. Heinosaari, and A. Toigo. Commutative POVMs and fuzzy observables. *Found. Phys.*, 39:593–612, 2009.
- [10] P. Lahti and S. Pulmannová. Coexistent observables and effects in quantum mechanics. *Rep. Math. Phys.*, 39:339–351, 1997.
- [11] Y. Kuramochi. Minimal sufficient positive-operator valued measure on a separable Hilbert space. *J. Math. Phys.*, 56:102205, 2015.
- [12] P. Stano, D. Reitzner, and T. Heinosaari. Coexistence of qubit effects. *Phys. Rev. A*, 78:012315, 2008.
- [13] P. Busch and H.-J. Schmidt. Coexistence of qubit effects. *Quantum Inf. Process.*, 9:143–169, 2010.
- [14] 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.
- [15] P. Busch and T. Heinosaari. Approximate joint measurements of qubit observables. *Quant. Inf. Comp.*, 8:0797–0818, 2008.
- [16] P. Busch. Unsharp reality and joint measurements for spin observables. *Phys. Rev. D*, 33:2253–2261, 1986.
- [17] M. Florenzano and C. Le Van. *Finite Dimensional Convexity and Optimization*. Springer, 2001.

APPENDIX A. ORDER THEORETIC DEFINITIONS

Let (X, \leq) be a partially ordered set (poset). Any subset Y of X is also a poset.

A poset X is *totally ordered* if for any $x, y \in X$, either $x \leq y$ or $y \leq x$ holds. A totally ordered subset C of X is called a *chain*.

- X is *upper directed* (or just *directed*) $:\stackrel{\text{def.}}{\Leftrightarrow}$ for any $x, y \in X$ there exists $z \in X$ such that $x \leq z$ and $y \leq z$.
- X is *lower directed* $:\stackrel{\text{def.}}{\Leftrightarrow}$ for any $x, y \in X$ there exists $z \in X$ such that $x \geq z$ and $y \geq z$.
- X is *upper* (resp. *lower*) *inductive* $:\stackrel{\text{def.}}{\Leftrightarrow}$ any chain of X has an upper bound (resp. a lower bound).
- X is *upper* (resp. *lower*) *directed complete* $:\stackrel{\text{def.}}{\Leftrightarrow}$ every upper (resp. lower) directed subset of X has a supremum (resp. an infimum).

It is immediate from the above definition that an upper (resp. a lower) directed complete poset is upper (resp. lower) inductive. According to Zorn's lemma, any upper (resp. lower) inductive poset has a maximal (resp. minimal) element.

Let (I, \leq) be a directed set and let X be a set. A map $I \ni i \mapsto x_i \in X$ is called a net on X . If X is a poset and for any $i, j \in I$, $i \leq j$ implies $x_i \leq x_j$ (resp. $x_j \leq x_i$), the net $(x_i)_{i \in I}$ is said to be monotonically increasing (resp. decreasing). A poset X is upper (resp. lower) directed complete if and only if the image of any monotonically increasing (resp. decreasing) net on X has a supremum (resp. an infimum).

APPENDIX B. PAIRWISE LINEARLY INDEPENDENT OBSERVABLES

An observable \mathbf{A} with an outcome set Ω is said to be *pairwise linearly independent* if any pair $(\mathbf{A}(x_1), \mathbf{A}(x_2))$, $x_1, x_2 \in \Omega$; $x_1 \neq x_2$, is linearly independent. Every observable is post-processing equivalent to a pairwise linearly independent observable unique up to the permutation of the outcome set [11]. An observable \mathbf{A} is pairwise linearly independent if and only if \mathbf{A} is minimal sufficient, that is, for any Markov kernel $p \in \mathbf{Markov}(\Omega, \Omega)$ the condition $p * \mathbf{A} = \mathbf{A}$ implies $p(x, x') = \delta_{x, x'}$ [11]. We recall that two Markov kernels $p \in \mathbf{Markov}(\Omega_1, \Omega_2)$ and $q \in \mathbf{Markov}(\Omega_2, \Omega_3)$ can be combined into a new Markov kernel as follows:

$$(p * q)(x, x') := \sum_{y \in \Omega_2} p(x, y)q(y, x').$$

The following proposition characterizes a Markov kernel that conserves the information of the post-processed observable.

Proposition 4. *Let \mathbf{A} be a pairwise linearly independent observable with an outcome set Ω and let $p \in \mathbf{Markov}(\Omega', \Omega)$ be a Markov kernel from Ω to a finite set Ω' . Then the following conditions are equivalent.*

- (i) $p * \mathbf{A} \sim_{\text{post}} \mathbf{A}$.
- (ii) For each $y \in \Omega'$, the finite set $\{p(y, x) : x \in \Omega\}$ has at most one nonzero element.

Proof. (i) \implies (ii). Assume (i). By assumption there exists a Markov kernel $q \in \mathbf{Markov}(\Omega, \Omega')$ such that $(q * p) * \mathbf{A} = q * (p * \mathbf{A}) = \mathbf{A}$. From the minimal sufficiency of \mathbf{A} follows that $(q * p)(x, x') = \delta_{x, x'}$. Now, assume that there exist $y_0 \in \Omega'$ and $x_1, x_2 \in \Omega$ with $x_1 \neq x_2$ such that $p(y_0, x_1)p(y_0, x_2) \neq 0$. We can take $x_0 \in \Omega$ such that $q(x_0, y_0) \neq 0$. Then for $i = 1, 2$,

$$\begin{aligned} \delta_{x_0, x_i} &= \sum_{y \in \Omega'} q(x_0, y)p(y, x_i) \\ &\geq q(x_0, y_0)p(y_0, x_i) \\ &> 0. \end{aligned}$$

This implies $x_1 = x_0 = x_2$, contradicting the assumption $x_1 \neq x_2$. Thus the condition (ii) holds.

(ii) \implies (i). Assume (ii). Let $\Omega'_1 := \{y \in \Omega' \mid (p * \mathbf{A})(y) \neq 0\}$ and let \mathbf{B} be an observable with the outcome set Ω'_1 , obtained by restricting $p * \mathbf{A}$ to Ω'_1 . Obviously, \mathbf{B} is post-processing equivalent to $p * \mathbf{A}$. By the assumption, for each $y \in \Omega'_1$ there exists a unique element $x_y \in \Omega$ such that $p(y, x_y) \neq 0$. Then $\mathbf{B}(y) = p(y, x_y)\mathbf{A}(x_y)$ and hence

$$\mathbf{A}(x) = \sum_{y \in \Omega'_1} \delta_{x, x_y} p(y, x_y) \mathbf{A}(x_y) = \sum_{y \in \Omega'_1} \delta_{x, x_y} \mathbf{B}(y),$$

which implies $\mathbf{A} \sim_{\text{post}} \mathbf{B} \sim_{\text{post}} p * \mathbf{A}$. \square

Corollary 4. *Let \mathbf{A} be an observable with an outcome set Ω and let $p \in \mathbf{Markov}(\Omega', \Omega)$ be a Markov kernel from Ω to a finite set Ω' . Then $p * \mathbf{A} \sim_{\text{post}} \mathbf{A}$ if and only if for each $y \in \Omega'$ and each $x_1, x_2 \in \Omega$, if $\mathbf{A}(x_1)$ and $\mathbf{A}(x_2)$ are linearly independent then $p(y, x_1)p(y, x_2) = 0$.*

Proof. We take a pairwise linearly independent observable $\mathbf{B}: \Omega_0 \rightarrow \mathcal{L}(\mathcal{H})$ post-processing equivalent to \mathbf{A} . Then there exist $q \in \mathbf{Markov}(\Omega, \Omega_0)$ and $y: \Omega \rightarrow \Omega_0$ such that $\mathbf{A}(x) = q(x, z(x))\mathbf{B}(z(x))$ ($x \in \Omega$), $q(x, z') = 0$ if $z(x) \neq z'$, and $\mathbf{A} = q * \mathbf{B}$. Hence $p * \mathbf{A} = (p * q) * \mathbf{B}$, and

$$p * q(y, z) = \sum_{x \in \Omega} p(y, x)q(x, z) = \sum_{x \in \Omega, z(x)=z} p(y, x)q(x, z) \quad (25)$$

Now assume $\mathbf{A} \sim_{\text{post}} p * \mathbf{A}$. Then $\mathbf{B} \sim_{\text{post}} (p * q) * \mathbf{B}$ and Proposition 4 implies that for each $y \in \Omega'$ $p * q(y, z) \neq 0$ holds at most one $z \in \Omega_0$. If $\mathbf{A}(x_1)$ and $\mathbf{A}(x_2)$ are linearly independent, then $z(x_1) \neq z(x_2)$ and $q(x_1, z(x_1)) \neq 0 \neq q(x_2, z(x_2))$. Hence (25) implies $p(y, x_1)p(y, x_2) = 0$. Conversely, if \mathbf{A} is not post-processing equivalent to $p * \mathbf{A}$, then from $p * \mathbf{A} \sim_{\text{post}} (p * q) * \mathbf{B}$ and Proposition 4, there exist $y \in \Omega'$ and $z_1, z_2 \in \Omega_0$ such that $p * q(y, z_1)p * q(y, z_2) \neq 0$ and $z_1 \neq z_2$. From (25) we can take $x_1, x_2 \in \Omega$ such that $z(x_1) = z_1$, $z(x_2) = z_2$, and $p(y, x_1)q(x_1, z_1)p(y, x_2)q(x_2, z_2) \neq 0$. Then $\mathbf{A}(x_1)$ and $\mathbf{A}(x_2)$ are linearly independent and $p(y, x_1)p(y, x_2) \neq 0$. \square

APPENDIX C. FINITE-DIMENSIONAL POLYHEDRA

This appendix briefly describes some facts about the finite-dimensional polyhedra used in the main part.

A convex set K on \mathbb{R}^n is called a *polyhedron* if there exist $m \geq 1$, $\mathbf{a}^i \in \mathbb{R}^n$, and $\alpha^i \in \mathbb{R}$ ($i = 1, \dots, m$) such that

$$K = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{a}^i \cdot \mathbf{x} \geq \alpha^i \ (\forall i \in \{1, \dots, m\})\},$$

i.e. K is the set of solutions of a finite number of linear inequalities. If a polyhedron K is compact, K is called a *polytope*. The set of extreme points $\text{ex}(K)$ of the polytope K is a finite set and can be calculated as follows. Consider a subset $I \subseteq \{1, \dots, m\}$ such that the system of linear equalities

$$\mathbf{a}^i \cdot \mathbf{x} = \alpha^i \quad (i \in I) \tag{26}$$

has a unique solution. Let \mathcal{I} be the set of such subsets of $\{1, \dots, m\}$ and let \mathbf{x}_I be the unique solution of (26) for each $I \in \mathcal{I}$. Then $\text{ex}(K)$ is given by

$$\text{ex}(K) = \{\mathbf{x}_I \mid I \in \mathcal{I}, \mathbf{x}_I \in K\}.$$

([17], Proposition 3.3.1).

A subset $C \subseteq \mathbb{R}^n$ is called a (pointed) *polyhedral cone* if there exist $m \geq 1$ and $\mathbf{a}^i \in \mathbb{R}^n$ ($i = 1, \dots, m$) such that

$$C = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{a}^i \cdot \mathbf{x} \geq 0, (\forall i \in \{1, \dots, m\})\},$$

i.e. C is the set of solutions of a finite number of *homogeneous* linear inequalities. For simplicity we assume that the linear span of $\{\mathbf{a}^i\}_{i=1}^m$ coincides with the full space \mathbb{R}^n . Then if $C \neq \{0\}$, C is a conical hull of some finite set A and A can be calculated as follows. Consider a subset $J \subseteq \{1, \dots, m\}$ such that the rank of the linear equalities

$$\mathbf{a}^i \cdot \mathbf{x} = 0 \quad (i \in J) \tag{27}$$

is $n - 1$ and let \mathcal{J} denote the set of such subsets. For each $J \in \mathcal{J}$ we denote by Δ_J the line consisting of the solutions of (27). Then

either $C \cap \Delta_J = \{0\}$ or $C \cap \Delta_J$ is a 1-dimensional face of C ([17], Proposition 3.3.2). Let $\mathcal{J}' := \{J \in \mathcal{J} \mid C \cap \Delta_J \neq \{0\}\}$. For each $J \in \mathcal{J}'$ we take $\mathbf{y}_J \in (C \cap \Delta_J) \setminus \{0\}$. Then A can be given by $A = \{\mathbf{y}_J\}_{J \in \mathcal{J}'}$. If $\mathcal{J}' = \emptyset$, we have $C = \{0\}$.