Infinite dimensionality of the post-processing order of measurements on a general state space

Yui Kuramochi

Department of Physics, Faculty of Science, Kyushu University, 744 Motooka, Nishi-ku, Fukuoka, Japan

E-mail: `yui.tasuke.kuramochi@gmail.com`

25 June 2022

Abstract. For a partially ordered set $(S, \preceq)$, the order (monotone) dimension is the minimum cardinality of total orders (respectively, real-valued order monotone functions) on S that characterize the order $\preceq$. In this paper we consider an arbitrary generalized probabilistic theory and the set of finite-outcome measurements on it, which can be described by effect-valued measures, equipped with the classical post-processing orders. We prove that the order and order monotone dimensions of the post-processing order are (countably) infinite if the state space is not a singleton (and is separable in the norm topology). This result gives a negative answer to the open question for quantum measurements posed in [Guff T *et al.* 2021 *J. Phys. A: Math. Theor.* **54** 225301]. We also consider the quantum post-processing relation of channels with a fixed input quantum system described by a separable Hilbert space $\mathcal{H}$ and show that the order (monotone) dimension is countably infinite when $\dim \mathcal{H} \geq 2$.

Keywords: effect-valued measure, quantum channels, post-processing relation of measurements, state discrimination probability, order dimension

1. Introduction

Order structures are ubiquitous in broad range of physical theories. Examples of them are the adiabatic accessibility relation in thermodynamics [1, 2], the convertibility relation by local operations and classical communication (LOCC) of bipartite quantum states [3, 4], the post-processing relation of quantum measurements, or positive operator-valued measures (POVMs) [5, 6, 7, 8, 9, 10, 11], and so on. These relations are related to (quantum) resource theories [12], which are recently being intensively studied in the area of quantum information.

In thermodynamics, the adiabatic accessibility relation is characterized by a single function, the entropy. This means that a thermodynamic state is adiabatically convertible to another one if and only if the entropy increases. From this, we can see that the adiabatic accessibility relation is a total order. On the other hand, the

Infinite dimensionality of the post-processing order of measurements   2

LOCC convertibility relation of finite-dimensional bipartite pure state is not a total order, while the well-known characterization of this relation by the majorization order ([4], Theorem 12.15) suggests that we have still finite number of order monotones (i.e. order-preserving real-valued functions) that characterize the order.

How about the case of the post-processing relation of POVMs? It is shown by many authors [13, 14, 11, 15] that the state discrimination probabilities characterize the post-processing order (Proposition 1), while they are uncountably infinite as pointed out in [11]. These in mind, the authors of [11] asked if we can characterize the post-processing order of POVMs by a *finite* number of order monotones, as in the case of the LOCC convertibility relation of bipartite pure states.

In this paper we give a negative answer to this open question. More strongly, we show that any choice of finite number of order monotones or total orders cannot characterize the post-processing relation of measurements on any non-trivial generalized probabilistic theory (GPT) [16, 17, 15], which is a general framework of physical theories and contains the quantum and classical theories as special examples. Moreover, we demonstrate that a *countable* number of order monotones characterizing the post-processing order exists if the state space is separable in the norm topology. As a corollary, we also prove a similar statement for the quantum post-processing relation for quantum channels with a fixed input Hilbert space.

The main theorems (Theorems 1 and 2) and their proofs are based on the notion of the *order dimension* (Definition 1) known in the area of order theory [18, 19, 20]. The order dimension of an order is defined as the minimum number of total orders that characterize the order and roughly quantifies the complexity of the order or its deviation from a simple total order.

In the main part, we also introduce another related quantity called the *order monotone dimension* of an order (Definition 1). This is defined as the minimum number of order monotones that characterize the order and directly connected to the open question in [11]. The order monotone dimension is shown to be always greater than or equal to the order dimension (see Lemma 2). The present work is the first attempt to evaluate these dimensions of orders appearing in quantum information.

The rest of this paper is organized as follows. In Section 2 we introduce some preliminaries and definitions. In Section 3 we state and prove the main theorems (Theorems 1 and 2). We then conclude the paper in Section 4.

2. Preliminaries

In this section, we give preliminaries on the order theory and post-processing relations of measurements and quantum channels and fix the notation.

In this paper, we denote by $\mathbb{N} = \{1, 2, \ldots\}$, $\mathbb{R}$, and $\mathbb{C}$ the sets of natural, real, and complex numbers, respectively. For each $m \in \mathbb{N}$, we write as $\mathbb{N}_m := \{0, 1, \ldots, m-1\}$. The vector spaces in this paper are over $\mathbb{R}$ unless otherwise stated.

2.1. Order theory

In this paper we identify each binary relation R on a set S with its graph, that is, R is the subset of $S \times S$ consisting of pairs satisfying the relation R. For a binary relation $R \subseteq S \times S$, the relation $(x, y) \in R$ $(x, y \in S)$ is occasionally written as xRy, which is consistent with the common notation. If A is a subset of S and R is a binary relation on S, we define the *restriction* of R to A by $R|_A := R \cap (A \times A)$.

Let R be a binary relation on a set S. We consider the following conditions for R.

- R is *reflexive* $:\overset{\text{def.}}{\Leftrightarrow} xRx$ $(\forall x \in S)$.
- R is *symmetric* $:\overset{\text{def.}}{\Leftrightarrow} xRy$ implies yRx $(\forall x, y \in S)$.
- R is *antisymmetric* $:\overset{\text{def.}}{\Leftrightarrow} xRy$ and yRx imply $x = y$ $(\forall x, y \in S)$.
- R is *transitive* $:\overset{\text{def.}}{\Leftrightarrow} xRy$ and yRz imply xRz $(\forall x, y, z \in S)$.
- R is *total* $:\overset{\text{def.}}{\Leftrightarrow}$ either xRy or yRx holds $(\forall x, y \in S)$.

A reflexive, transitive binary relation is called a *preorder*. An antisymmetric preorder is called a *partial order*. A total partial order is just called a *total order* (or a *linear order*). If R is a partial order on a set S, (S, R) is called a *partially ordered set*, or a *poset*. A poset (S, R) is called a *chain* if R is a total order.

If P is a preorder on S, then the relation $\sim$ defined via

$$x \sim y :\overset{\text{def.}}{\Leftrightarrow} xPy \text{ and } yPx \quad (x, y \in S)$$

is an equivalence relation (i.e. a reflexive, symmetric, and transitive relation). If we denote by $[x]$ the equivalence class to which $x \in S$ belongs, then we may define a binary relation R on the quotient space $S/\sim$ via

$$[x]R[y] :\overset{\text{def.}}{\Leftrightarrow} xPy$$

and R is a partial order on $S/\sim$.

A binary relation P on S is said to be an *extension* of a binary relation R on S if $R \subseteq P$, i.e. xRy implies xPy for every $x, y \in S$. If L is an extension of a binary relation R and is a total order, then L is called a *linear extension* of R.

A family $\mathcal{L}$ of binary relations on a set S is said to *realize* a binary relation R on S (or *realize* (S, R)) if $R = \bigcap \mathcal{L}$, i.e. for every $x, y \in S$

$$xRy \iff [xLy \quad (\forall L \in \mathcal{L})].$$

Let $(S, \preceq)$ be a poset. A real-valued function $f \colon S \to \mathbb{R}$ is called an *order monotone*, or a *resource monotone*, (with respect to $\preceq$) if for every $x, y \in S$

$$x \preceq y \implies f(x) \le f(y).$$

A family $\mathcal{F}$ of real-valued functions on S is said to *characterize* (or to be a *complete family* of) $(S, \preceq)$ if for every $x, y \in S$

$$x \preceq y \iff [f(x) \le f(y) \quad (\forall f \in \mathcal{F})]. \tag{1}$$

Infinite dimensionality of the post-processing order of measurements 4

If (1) holds, then each $f \in \mathcal{F}$ is necessarily an order monotone.

The following notions of the order and order monotone dimensions play the central role in this paper.

Definition 1. Let $(S, \preceq)$ be a poset.

(i) By the *order dimension* [18, 19] of $(S, \preceq)$, written as $\dim_{\mathrm{ord}}(S, \preceq)$, we mean the minimum cardinality [21] $|\mathcal{L}|$ of a family $\mathcal{L}$ of linear extensions of $\preceq$ that realizes $(S, \preceq)$. Here $|A|$ denotes the cardinality of a set A.

(ii) By the *order monotone dimension* of $(S, \preceq)$, written as $\dim_{\mathrm{ord}, \mathbb{R}}(S, \preceq)$, we mean the minimum cardinality $|\mathcal{F}|$ of a family $\mathcal{F}$ of order monotones on S that characterizes $(S, \preceq)$.

The order monotone dimension of a poset $(S, \preceq)$ is well-defined. Indeed, according to [22] (Proposition 5.2), if we define the order monotone

$$M_a(x) := \begin{cases} 1 & \text{(when } a \preceq x\text{)}; \\ 0 & \text{(otherwise)} \end{cases} \quad (x \in S)$$

for each $a \in S$, then the family $\{M_a\}_{a \in S}$ characterizes $(S, \preceq)$. Hence $\dim_{\mathrm{ord}, \mathbb{R}}(S, \preceq)$ is well-defined and at most $|S|$. The well-definedness of the order dimension is proved in [18] (Theorem 2.32) by using the Szpilrajn extension theorem [23]. We also note that, since the cardinals are well-ordered [21], we can always take a family $\mathcal{F}$ of monotones that characterizes $\preceq$ and the cardinality $|\mathcal{F}|$ is minimum, i.e. $|\mathcal{F}| = \dim_{\mathrm{ord}, \mathbb{R}}(S, \preceq)$. A similar statement for the order dimension is also true. We will prove in Lemma 2 that the order monotone dimension is always greater than or equal to the order dimension.

2.2. General probabilistic theory

A general probabilistic theory (GPT) with the no-restriction hypothesis [24] is mathematically described by the following notion of the base-norm Banach space.

Definition 2. A triple (V, V_+, Ω) is called a *base-norm Banach space* if the following conditions hold.

(i) V is a real vector space.

(ii) V_+ is a positive cone of V, i.e. $\lambda V_+ \subseteq V_+$ ($\forall \lambda \in [0, \infty)$), $V_+ + V_+ \subseteq V_+$, and $V_+ \cap (-V_+) = \{0\}$ hold. We define the linear order on V induced from V_+ by

$$x \leq y :\overset{\text{def.}}{\Leftrightarrow} y - x \in V_+ \quad (x, y \in V).$$

(iii) V_+ is generating, i.e. $V = V_+ + (-V_+)$.

(iv) Ω is a base of V_+, i.e. Ω is a convex subset of V_+ and for every $x \in V_+$ there exists a unique $\lambda \in [0, \infty)$ such that $x \in \lambda \Omega$.

Infinite dimensionality of the post-processing order of measurements 5

(v) We define the base-norm on V by

$$\|x\| := \inf \left\{ \alpha + \beta \mid x = \alpha\omega_1 + \beta\omega_2;\ \alpha, \beta \in [0, \infty);\ \omega_1, \omega_2 \in \Omega \right\} \quad (x \in V).$$

We require that the base-norm $\|\cdot\|$ is a complete norm on V.

If these conditions are satisfied, Ω is called a *state space* and each element $\omega \in \Omega$ is called a *state*.

The reader can find in [17] (Chapter 1) how the notion of the base-norm Banach space is derived from operationally natural requirements on the GPT.

Let (V, V_+, Ω) be a base-norm Banach space. We denote by V^* the continuous dual of V (i.e. the set of norm-continuous real linear functionals on V) equipped with the dual norm

$$\|f\| := \sup_{x \in V, \|x\| \leq 1} |f(x)| \quad (f \in V^*).$$

For each $f \in V^*$ and each $x \in V$ we occasionally write as $\langle f, x \rangle := f(x)$. The dual positive cone V_+^* of V^* is defined by

$$V_+^* := \left\{ f \in V^* \mid \langle f, x \rangle \geq 0\, (\forall x \in V_+) \right\}$$

and the dual linear order by

$$f \leq g :\overset{\text{def.}}{\Leftrightarrow} g - f \in V_+^* \Leftrightarrow [f(x) \leq g(x) \quad (\forall x \in V_+)] \qquad (f, g \in V^*).$$

It can be shown that there exists a unique positive element, called the unit element, $u_\Omega \in V_+^*$ such that $\langle u_\Omega, \Omega \rangle = 1$. Then the dual norm on V^* coincides with the order unit norm of u_Ω:

$$\|f\| = \inf \left\{ \lambda \in [0, \infty) \mid -\lambda u \leq f \leq \lambda u \right\} \quad (f \in V^*).$$

An element $e \in V^*$ satisfying $0 \leq e \leq u_\Omega$ is called an *effect* (on Ω). The set of effects on Ω is denoted by $\mathcal{E}(\Omega)$.

In the main part of the paper, we will consider the following examples of quantum and classical theories.

Example 1 (Quantum theory)**.** Let $\mathcal{H}$ be a complex Hilbert space. We write the inner product of $\mathcal{H}$ as $\langle \cdot | \cdot \rangle$ which is antilinear and linear in the first and second components, respectively, and the complete norm as $\|\psi\| := \langle \psi | \psi \rangle^{1/2}$. The sets of bounded and trace-class linear operators on $\mathcal{H}$ are denoted by $\mathbf{B}(\mathcal{H})$ and $\mathbf{T}(\mathcal{H})$, respectively. The self-adjoint and positive parts of these sets are defined by

$$\mathbf{B}_{\mathrm{sa}}(\mathcal{H}) := \left\{ a \in \mathbf{B}(\mathcal{H}) \mid a = a^* \right\},$$
$$\mathbf{B}_+(\mathcal{H}) := \left\{ a \in \mathbf{B}(\mathcal{H}) \mid \langle \psi | a\psi \rangle \geq 0\, (\forall \psi \in \mathcal{H}) \right\},$$
$$\mathbf{T}_{\mathrm{sa}}(\mathcal{H}) := \left\{ a \in \mathbf{T}(\mathcal{H}) \mid a = a^* \right\},$$
$$\mathbf{T}_+(\mathcal{H}) := \left\{ a \in \mathbf{T}(\mathcal{H}) \mid \langle \psi | a\psi \rangle \geq 0\, (\forall \psi \in \mathcal{H}) \right\},$$

Infinite dimensionality of the post-processing order of measurements   6

where a^* denotes the adjoint operator of $a \in \mathbf{B}(\mathcal{H})$. The uniform and the trace norms are respectively defined by

$$\|a\| := \sup_{\psi \in \mathcal{H}, \|\psi\| \leq 1} \|a\psi\| \quad (a \in \mathbf{B}(\mathcal{H})),$$

$$\|b\|_1 := \mathrm{tr}(\sqrt{b^* b}) \quad (b \in \mathbf{T}(\mathcal{H})),$$

where $\mathrm{tr}(\cdot)$ denotes the trace. A non-negative trace-class operator ρ satisfying the normalization condition $\mathrm{tr}(\rho) = 1$ is called a density operator. The set of density operators on $\mathcal{H}$ is denoted by $\mathbf{D}(\mathcal{H})$.

In the GPT framework in Definition 2, the quantum theory corresponds to the case $(V, V_+, \Omega) = (\mathbf{T}_{\mathrm{sa}}(\mathcal{H}), \mathbf{T}_+(\mathcal{H}), \mathbf{D}(\mathcal{H}))$. The base-norm on $\mathbf{T}_{\mathrm{sa}}(\mathcal{H})$ then coincides with the trace norm. The continuous dual space $\mathbf{T}_{\mathrm{sa}}(\mathcal{H})^*$, the dual norm on it, the dual positive cone $\mathbf{T}_{\mathrm{sa}}(\mathcal{H})^*_+$, and the unit element $u_{\mathbf{D}(\mathcal{H})}$ are respectively identified with $\mathbf{B}_{\mathrm{sa}}(\mathcal{H})$, the uniform norm, $\mathbf{B}_+(\mathcal{H})$, and the identity operator $\mathbb{1}_{\mathcal{H}}$ on $\mathcal{H}$ by the duality

$$\langle a, b \rangle = \mathrm{tr}(ab) \quad (a \in \mathbf{B}_{\mathrm{sa}}(\mathcal{H}); \, b \in \mathbf{T}_{\mathrm{sa}}(\mathcal{H})).$$

By this duality, we identify $\mathbf{T}_{\mathrm{sa}}(\mathcal{H})^*$ with $\mathbf{B}_{\mathrm{sa}}(\mathcal{H})$.

Example 2 (Discrete classical theory)**.** Let X be a non-empty set. We define $\ell^1(X)$ and $\ell^\infty(X)$ and their positive parts by

$$\ell^1(X) := \{\, f = (f(x))_{x \in X} \in \mathbb{R}^X \mid \|f\|_1 < \infty \,\},$$
$$\ell^\infty(X) := \{\, f = (f(x))_{x \in X} \in \mathbb{R}^X \mid \|f\|_\infty < \infty \,\},$$
$$\ell^1_+(X) := \{\, (f(x))_{x \in X} \in \ell^1(X) \mid f(x) \geq 0 \, (\forall x \in X) \,\},$$
$$\ell^\infty_+(X) := \{\, (f(x))_{x \in X} \in \ell^\infty(X) \mid f(x) \geq 0 \, (\forall x \in X) \,\},$$

where for $f \in \mathbb{R}^X$

$$\|f\|_1 := \sum_{x \in X} |f(x)|,$$

$$\|f\|_\infty := \sup_{x \in X} |f(x)|$$

are respectively the ℓ^1- and the ℓ^∞-norms. We also define the simplex of the probability distributions on X by

$$\mathcal{P}(X) := \left\{\, (p(x))_{x \in X} \in \ell^1_+(X) \,\middle|\, \sum_{x \in X} p(x) = 1 \,\right\}.$$

In the framework of the GPT in Definition 2, a discrete classical theory corresponds to the case $(V, V_+, \Omega) = (\ell^1(X), \ell^1_+(X), \mathcal{P}(X))$. The base-norm on $\ell^1(X)$ coincides with the ℓ^1-norm $\|\cdot\|_1$. The continuous dual $\ell^1(X)^*$, the dual norm on $\ell^1(X)^*$, the dual positive cone $\ell^1(X)^*_+$, and the unit element $u_{\mathcal{P}(X)}$ are respectively identified with $\ell^\infty(X)$, the ℓ^∞-norm $\|\cdot\|_\infty$, $\ell^\infty_+(X)$, and the constant function $1_X := (1)_{x \in X}$ by the duality

$$\langle f, g \rangle = \sum_{x \in X} f(x)g(x) \quad (f \in \ell^\infty(X); \, g \in \ell^1(X)).$$

Infinite dimensionality of the post-processing order of measurements   7

2.3. Post-processing relation of measurements on a GPT

Now we fix a base-norm Banach space (V, V_+, Ω) corresponding to a GPT.

For a natural number $m \in \mathbb{N}$, a finite sequence $(\mathsf{M}(k))_{k=0}^{m-1} \in (V^*)^m$ is called an *(m-outcome) effect-valued measure* (EVM) (on Ω) if $\mathsf{M}(k) \geq 0$ $(k \in \mathbb{N}_m)$ and $\sum_{k=0}^{m-1} \mathsf{M}(k) = u_\Omega$ hold. We denote by $\mathrm{EVM}_m(\Omega)$ the set of m-outcome EVMs. We also define $\mathrm{EVM}_{\mathrm{fin}}(\Omega) := \bigcup_{m \in \mathbb{N}} \mathrm{EVM}_m(\Omega)$, which is the set of finite-outcome EVMs on Ω. For each EVM $\mathsf{M} = (\mathsf{M}(k))_{k=0}^{m-1}$ and each state $\omega \in \Omega$, the sequence $(p_\omega^\mathsf{M}(k))_{k=0}^{m-1}$ defined by

$$p_\omega^\mathsf{M}(k) := \langle \mathsf{M}(k), \omega \rangle \quad (k \in \mathbb{N}_m)$$

is a probability distribution. In the physical context, p_ω^M is the outcome probability distribution of the measurement M when the state of the system is prepared to be ω.

Let $\mathsf{M} = (\mathsf{M}(j))_{j=0}^{m-1}$ and $\mathsf{N} = (\mathsf{N}(k))_{k=0}^{n-1}$ be EVMs on Ω. M is said to be a *post-processing* of (or, *less or equally informative* than) N [5, 6, 7, 8, 9, 10], written as $\mathsf{M} \preceq_{\mathrm{post}} \mathsf{N}$, if there exists a matrix $(p(j|k))_{j \in \mathbb{N}_m, k \in \mathbb{N}_n}$ such that

$$p(j|k) \geq 0 \quad (j \in \mathbb{N}_m,\ k \in \mathbb{N}_n), \tag{2}$$

$$\sum_{j=0}^{m-1} p(j|k) = 1 \quad (k \in \mathbb{N}_n), \tag{3}$$

$$\mathsf{M}(j) = \sum_{k=0}^{n-1} p(j|k)\mathsf{N}(k) \quad (j \in \mathbb{N}_m).$$

A matrix $(p(j|k))_{j \in \mathbb{N}_m, k \in \mathbb{N}_n}$ satisfying (2) and (3) is called a *Markov matrix*. The relation $\mathsf{M} \preceq_{\mathrm{post}} \mathsf{N}$ means that the measurement M is realized if we first perform N, which gives a measurement outcome $k \in \mathbb{N}_n$, then randomly generate $j \in \mathbb{N}_m$ according to the probability distribution $(p(j|k))_{j \in \mathbb{N}_m}$, forget the original measurement outcome k, and finally record j as the measurement outcome. We also say that M and N are *post-processing equivalent* (or *equally informative*), written as $\mathsf{M} \sim_{\mathrm{post}} \mathsf{N}$, if both $\mathsf{M} \preceq_{\mathrm{post}} \mathsf{N}$ and $\mathsf{N} \preceq_{\mathrm{post}} \mathsf{M}$ hold. The binary relations $\preceq_{\mathrm{post}}$ and $\sim_{\mathrm{post}}$ are respectively preorder and equivalence relations on $\mathrm{EVM}_{\mathrm{fin}}(\Omega)$.

We write as $\mathfrak{M}_{\mathrm{fin}}(\Omega) := \mathrm{EVM}_{\mathrm{fin}}(\Omega)/\sim_{\mathrm{post}}$ and, for each $\mathsf{M} \in \mathrm{EVM}_{\mathrm{fin}}(\Omega)$, denote by $[\mathsf{M}]$ the equivalence class to which M belongs. We define the binary relation $\preceq_{\mathrm{post}}$ on $\mathfrak{M}_{\mathrm{fin}}(\Omega)$ by

$$[\mathsf{M}] \preceq_{\mathrm{post}} [\mathsf{N}] :\overset{\mathrm{def.}}{\Leftrightarrow} \mathsf{M} \preceq_{\mathrm{post}} \mathsf{N} \quad ([\mathsf{M}], [\mathsf{N}] \in \mathfrak{M}_{\mathrm{fin}}(\Omega)).$$

Then $(\mathfrak{M}_{\mathrm{fin}}(\Omega), \preceq_{\mathrm{post}})$ is a poset.

An EVM $\mathsf{M} = (\mathsf{M}(k))_{k=0}^{m-1}$ is called *trivial* if each component $\mathsf{M}(k)$ is proportional to the unit u_Ω. The equivalence class $[\mathsf{M}]$ of a trivial EVM M is the minimum element of the poset $(\mathfrak{M}_{\mathrm{fin}}(\Omega), \preceq_{\mathrm{post}})$, i.e. $[\mathsf{M}] \preceq_{\mathrm{post}} [\mathsf{N}]$ for all $[\mathsf{N}] \in \mathfrak{M}_{\mathrm{fin}}(\Omega)$.

The post-processing relation $\preceq_{\mathrm{post}}$ on $\mathfrak{M}_{\mathrm{fin}}(\Omega)$ is characterized by the state discrimination probabilities defined as follows. A finite sequence $\mathcal{E} = (\rho_k)_{k=0}^{N-1} \in V_+^N$ of non-negative elements in V is called an *ensemble* (on Ω) if the normalization condition

$\sum_{k=0}^{N-1} \langle u_\Omega, \rho_k \rangle = 1$ holds. For an ensemble $\mathcal{E} = (\rho_k)_{k=0}^{N-1}$ and an EVM $\mathsf{M} = (\mathsf{M}(j))_{j=0}^{m-1}$ on Ω, we define the *state discrimination probability* by

$$P_{\mathrm{g}}(\mathcal{E}; \mathsf{M}) := \sup_{(p(k|j))_{k\in\mathbb{N}_N,\, j\in\mathbb{N}_m}:\text{ Markov matrix}} \sum_{k\in\mathbb{N}_N,\, j\in\mathbb{N}_m} p(k|j) \langle \mathsf{M}(j), \rho_k \rangle \qquad (4)$$

$$= \sup_{\mathsf{A}\in\mathrm{EVM}(N;\mathsf{M})} \sum_{k\in\mathbb{N}_N} \langle \mathsf{A}(k), \rho_k \rangle,$$

where

$$\mathrm{EVM}(N; \mathsf{M}) := \left\{ \mathsf{A} = (\mathsf{A}(k))_{k=0}^{N-1} \in \mathrm{EVM}_N(\Omega) \mid \mathsf{A} \preceq_{\mathrm{post}} \mathsf{M} \right\},$$

is the set of N-outcome EVMs obtained by post-processing M. The ensemble $\mathcal{E} = (\rho_k)_{k=0}^{N-1}$ corresponds to the situation in which the system is prepared in the state $\langle u_\Omega, \rho_k \rangle^{-1} \rho_k$ with the probability $\langle u_\Omega, \rho_k \rangle$. The state discrimination probability (4) is the optimum average probability of the correct guessing of the index k of the ensemble when we are given the measurement outcome j of M. The maximum of the optimization problem in the RHS of (4) is attained when $p(k|j) = \delta_{k,k(j)}$, where $k(j)$ is chosen so that $k(j) \in \arg\max_{k\in\mathbb{N}_N} \langle \mathsf{M}(j), \rho_k \rangle$, i.e. when the maximum likelihood estimation is adopted. The optimal value is then given by

$$P_{\mathrm{g}}(\mathcal{E}; \mathsf{M}) = \sum_{j\in\mathbb{N}_m} \max_{k\in\mathbb{N}_N} \langle \mathsf{M}(j), \rho_k \rangle \qquad (5)$$

([15], Lemma 5).

The following Blackwell-Sherman-Stein (BSS) theorem for EVMs [13, 14, 11, 15] can be regarded as the generalization of the corresponding BSS theorem for statistical experiments known in mathematical statistics [25, 26].

Proposition 1 (BSS theorem for EVMs)**.** *Let* (V, V_+, Ω) *be a base-norm Banach space and let* M *and* N *be finite-outcome EVMs on* Ω. *Then* $\mathsf{M} \preceq_{\mathrm{post}} \mathsf{N}$ *if and only if* $P_{\mathrm{g}}(\mathcal{E}; \mathsf{M}) \le P_{\mathrm{g}}(\mathcal{E}; \mathsf{N})$ *for all ensemble* $\mathcal{E}$ *on* Ω.

The BSS theorem for continuous EVMs on a possibly infinite-dimensional GPT is proved in [15]. We give a straightforward proof of Proposition 1 in Appendix A. Proposition 1 implies that the functions

$$\mathfrak{M}_{\mathrm{fin}}(\mathcal{H}) \ni [\mathsf{M}] \mapsto P_{\mathrm{g}}(\mathcal{E}; \mathsf{M}) =: P_{\mathrm{g}}(\mathcal{E}; [\mathsf{M}]) \in \mathbb{R} \quad (\mathcal{E}\text{: ensemble})$$

are well-defined and the family

$$\left\{ P_{\mathrm{g}}(\mathcal{E}; \cdot) \mid \mathcal{E} \text{ is an ensemble on } \Omega \right\}$$

characterizes the poset $(\mathfrak{M}_{\mathrm{fin}}(\Omega), \preceq_{\mathrm{post}})$.

2.4. Post-processing relation of quantum channels

Let $\mathcal{H}$ and $\mathcal{K}$ be complex Hilbert spaces and let $\Gamma: \mathbf{B}(\mathcal{K}) \to \mathbf{B}(\mathcal{H})$ be a complex linear map. We have the following definitions.

- Γ is *unital* $:\overset{\text{def.}}{\Leftrightarrow} \Gamma(\mathbb{1}_{\mathcal{K}}) = \mathbb{1}_{\mathcal{H}}$.

- Γ is *positive* $:\overset{\text{def.}}{\Leftrightarrow} a \geq 0$ implies $\Gamma(a) \geq 0$ for every $a \in \mathbf{B}(\mathcal{K})$.

- Γ is *completely positive* (*CP*) [27, 28, 29] $:\overset{\text{def.}}{\Leftrightarrow}$ the product linear map $\Gamma \otimes \text{id}_n : \mathbf{B}(\mathcal{K} \otimes \mathbb{C}^n) \to \mathbf{B}(\mathcal{H} \otimes \mathbb{C}^n)$ is positive for all $n \in \mathbb{N}$, where $\text{id}_n : \mathbf{B}(\mathbb{C}^n) \to \mathbf{B}(\mathbb{C}^n)$ is the identity map on $\mathbf{B}(\mathbb{C}^n)$.

- Γ is a *channel* (in the Heisenberg picture) $:\overset{\text{def.}}{\Leftrightarrow} \Gamma$ is unital and CP.

- For positive Γ, Γ is *normal* $:\overset{\text{def.}}{\Leftrightarrow} \sup_i \Gamma(a_i) = \Gamma(\sup_i a_i)$ for every upper bounded increasing net (a_i) in $\mathbf{B}(\mathcal{K})$. This condition is equivalent to the ultraweak continuity of Γ, which means that $b_j \overset{\text{uw}}{\longrightarrow} b$ implies $\Gamma(b_j) \overset{\text{uw}}{\longrightarrow} \Gamma(b)$ for every net (b_j) and every element b in $\mathbf{B}(\mathcal{K})$, where $\overset{\text{uw}}{\longrightarrow}$ denotes the ultraweak (or σ-weak [29]) convergence.

For a channel $\Gamma : \mathbf{B}(\mathcal{K}) \to \mathbf{B}(\mathcal{H})$, the Hilbert spaces $\mathcal{H}$ and $\mathcal{K}$ are called respectively the input and output Hilbert spaces of Γ.

If $\Gamma : \mathbf{B}(\mathcal{K}) \to \mathbf{B}(\mathcal{H})$ is positive and normal, then there exists the unique positive linear map $\Gamma_* : \mathbf{T}(\mathcal{H}) \to \mathbf{T}(\mathcal{K})$, called the predual map of Γ, such that

$$\text{tr}(\rho \Gamma(a)) = \text{tr}(\Gamma_*(\rho)a) \quad (\rho \in \mathbf{T}(\mathcal{H}); a \in \mathbf{B}(\mathcal{K})). \tag{6}$$

Conversely if $\Gamma_* : \mathbf{T}(\mathcal{H}) \to \mathbf{T}(\mathcal{K})$ is a positive linear map then there exists unique normal positive linear map Γ satisfying (6). The predual of a normal channel describes the state change of the system (channel in the Schrödinger picture).

Let $\Gamma : \mathbf{B}(\mathcal{K}) \to \mathbf{B}(\mathcal{H})$ and $\Lambda : \mathbf{B}(\mathcal{J}) \to \mathbf{B}(\mathcal{H})$ be normal channels with the same input Hilbert space $\mathcal{H}$. We define the post-processing relations for channels as follows.

- $\Gamma \preceq_{\text{CP}} \Lambda$ (Γ is less or equally informative than Λ) $:\overset{\text{def.}}{\Leftrightarrow}$ there exists a normal channel $\Psi : \mathbf{B}(\mathcal{K}) \to \mathbf{B}(\mathcal{J})$ such that $\Gamma = \Lambda \circ \Psi$. It is known that $\Gamma \preceq_{\text{CP}} \Lambda$ holds if and only if there exists a (not necessarily normal) channel $\Phi : \mathbf{B}(\mathcal{K}) \to \mathbf{B}(\mathcal{J})$ such that $\Gamma = \Lambda \circ \Phi$ ([30], Lemma 3.12; [31], Theorem 2).

- $\Gamma \sim_{\text{CP}} \Lambda$ (Γ and Λ are equally informative) $:\overset{\text{def.}}{\Leftrightarrow} \Gamma \preceq_{\text{CP}} \Lambda$ and $\Lambda \preceq_{\text{CP}} \Gamma$.

The binary relations $\preceq_{\text{CP}}$ and $\sim_{\text{CP}}$ are respectively a preorder and an equivalence relation on the class $\mathbf{Ch}(\to \mathbf{B}(\mathcal{H}))$ of normal channels with a fixed input Hilbert space $\mathcal{H}$. It can be shown that there exist a set $\mathfrak{C}(\mathcal{H})$ and a class-to-set surjection

$$\mathbf{Ch}(\to \mathbf{B}(\mathcal{H})) \ni \Gamma \mapsto [\Gamma] \in \mathfrak{C}(\mathcal{H}) \tag{7}$$

such that

$$\Gamma \sim_{\text{CP}} \Lambda \iff [\Gamma] = [\Lambda]$$

for every $\Gamma, \Lambda \in \mathbf{Ch}(\to \mathbf{B}(\mathcal{H}))$. This follows from a more general result for normal channels with arbitrary input and output von Neumann algebras ([32], Section 3.3). We fix such a set $\mathfrak{C}(\mathcal{H})$ and a map (7). We also define the post-processing relation on $\mathfrak{C}(\mathcal{H})$ by

$$[\Gamma] \preceq_{\text{CP}} [\Lambda] :\overset{\text{def.}}{\Leftrightarrow} \Gamma \preceq_{\text{CP}} \Lambda \quad ([\Gamma], [\Lambda] \in \mathfrak{C}(\mathcal{H})).$$

Infinite dimensionality of the post-processing order of measurements 10

Then $(\mathfrak{C}(\mathcal{H}), \preceq_{\mathrm{CP}})$ is a poset.

Let $\mathcal{H}$ be a complex Hilbert space, let $\mathcal{E} = (\rho_k)_{k=0}^{n-1}$ be an ensemble on $\mathbf{D}(\mathcal{H})$, and let $\Gamma \colon \mathbf{B}(\mathcal{K}) \to \mathbf{B}(\mathcal{H})$ a channel (or more generally, a unital positive map). We define the state discrimination probability by

$$P_{\mathrm{g}}(\mathcal{E}; \Gamma) := \sup_{(\mathsf{M}(k))_{k=0}^{n-1} \in \mathrm{EVM}_n(\mathbf{D}(\mathcal{K}))} \sum_{k=0}^{n-1} \langle \Gamma(\mathsf{M}(k)), \rho_k \rangle. \tag{8}$$

This quantity is the maximal state discrimination probability of the index k of the ensemble $\mathcal{E}$ when the operation Γ is performed on the system whose state is prepared according to $\mathcal{E}$, and then an optimal measurement M on the output space $\mathcal{K}$ is performed.

The post-processing relation for channels is characterized by the state discrimination probabilities *with quantum side information* as shown in the following BSS-type theorem.

Proposition 2 (BSS theorem for channels). *Let $\Gamma \colon \mathbf{B}(\mathcal{K}) \to \mathbf{B}(\mathcal{H})$ and $\Lambda \colon \mathbf{B}(\mathcal{J}) \to \mathbf{B}(\mathcal{H})$ be normal channels. Then the following conditions are equivalent.*

(i) $\Gamma \preceq_{\mathrm{CP}} \Lambda$.

(ii) For every $n \in \mathbb{N}$ and every ensemble $\mathcal{E}$ on $\mathbf{D}(\mathcal{H} \otimes \mathbb{C}^n)$

$$P_{\mathrm{g}}(\mathcal{E}; \Gamma \otimes \mathrm{id}_n) \leq P_{\mathrm{g}}(\mathcal{E}; \Lambda \otimes \mathrm{id}_n) \tag{9}$$

holds.

Proposition 2 for finite-dimensional channels is proved in [33] by using Shmaya's theorem [34]. In Appendix B, we give a proof of Proposition 2 based on another (infinite-dimensional) BSS theorem for normal positive maps obtained in [35].

Proposition 2 implies that the function

$$\mathfrak{C}(\mathcal{H}) \ni [\Gamma] \mapsto P_{\mathrm{g}}(\mathcal{E}; \Gamma \otimes \mathrm{id}_n) =: P_{\mathrm{g}}^{(n)}(\mathcal{E}; [\Gamma]) \in \mathbb{R}$$

is a well-defined order monotone for every $n \in \mathbb{N}$ and every ensemble $\mathcal{E}$ on $\mathbf{D}(\mathcal{H} \otimes \mathbb{C}^n)$ and that the family

$$\{P_{\mathrm{g}}^{(n)}(\mathcal{E}; \cdot) \mid n \in \mathbb{N} \text{ and } \mathcal{E} \text{ is an ensemble on } \mathbf{D}(\mathcal{H} \otimes \mathbb{C}^n)\}$$

characterizes the poset $(\mathfrak{C}(\mathcal{H}), \preceq_{\mathrm{CP}})$.

3. Main theorems and their proofs

In this section we prove the following main theorems of this paper:

Theorem 1. *Let (V, V_+, Ω) be a base-norm Banach space with $\dim V \geq 2$. Then the following assertions hold.*

1. Both $\dim_{\mathrm{ord}}(\mathfrak{M}_{\mathrm{fin}}(\Omega), \preceq_{\mathrm{post}})$ and $\dim_{\mathrm{ord}, \mathbb{R}}(\mathfrak{M}_{\mathrm{fin}}(\Omega), \preceq_{\mathrm{post}})$ are infinite.

Infinite dimensionality of the post-processing order of measurements 11

2. *If V is separable in the norm topology, i.e. V has a countable norm dense subset, then*

$$\dim_{\mathrm{ord}}(\mathfrak{M}_{\mathrm{fin}}(\Omega), \preceq_{\mathrm{post}}) = \dim_{\mathrm{ord}, \mathbb{R}}(\mathfrak{M}_{\mathrm{fin}}(\Omega), \preceq_{\mathrm{post}}) = \aleph_0, \tag{10}$$

where $\aleph_0 = |\mathbb{N}|$ denotes the cardinality of a countably infinite set.

Theorem 2. *Let $\mathcal{H}$ be a separable complex Hilbert space with $\dim \mathcal{H} \geq 2$. Then*

$$\dim_{\mathrm{ord}}(\mathfrak{M}_{\mathrm{fin}}(\mathbf{D}(\mathcal{H})), \preceq_{\mathrm{post}}) = \dim_{\mathrm{ord}, \mathbb{R}}(\mathfrak{M}_{\mathrm{fin}}(\mathbf{D}(\mathcal{H})), \preceq_{\mathrm{post}}) = \aleph_0, \tag{11}$$

$$\dim_{\mathrm{ord}}(\mathfrak{C}(\mathcal{H}), \preceq_{\mathrm{CP}}) = \dim_{\mathrm{ord}, \mathbb{R}}(\mathfrak{C}(\mathcal{H}), \preceq_{\mathrm{CP}}) = \aleph_0. \tag{12}$$

Remark 1. The assumption $\dim V \geq 2$ in Theorem 1 holds if and only if the state space Ω contains at least two distinct points because V is the linear span of Ω and Ω is contained in the hyperplane $\{ x \in V \mid \langle u_\Omega, x \rangle = 1 \}$, which does not contain the origin. Similarly the separability of V in Theorem 1 is equivalent to that of Ω.

Remark 2. Theorem 1.1 is proved by explicitly constructing a sequence of finite subsets of $\mathfrak{M}_{\mathrm{fin}}(\Omega)$ with arbitrarily large dimensions (see Lemmas 3 and 7). Indeed, it is proved in [36] that for every poset with an infinite order dimension we can always find such a sequence of finite subsets. We give another simple proof of this fact in Appendix C based on Tychonoff's theorem.

Remark 3. As shown in [15], we can define the set $\mathfrak{M}(\Omega)$ of post-processing equivalence classes of general (i.e. possibly continuous outcome) EVMs on Ω. Then we can easily show that Theorem 1.1 is also valid for $\mathfrak{M}(\Omega)$ since $\mathfrak{M}_{\mathrm{fin}}(\Omega)$ is a subset of $\mathfrak{M}(\Omega)$, i.e. the finite-outcome EVM is a special kind of the general EVM. We can also show that the proof of Theorem 1.2 (Lemma 8) can be straightforwardly generalized to $\mathfrak{M}(\Omega)$ since the BSS theorem is also valid for general EVMs ([15], Theorem 1).

Remark 4. A parametrized family $(P_\theta)_{\theta \in \Theta}$ of classical probabilities with a fixed sample space is called a statistical experiment, or a statistical model, and is one of the basic concepts in mathematical statistics [25, 26]. As shown in [15] (Appendix D), we can identify the class of statistical experiments with a fixed parameter set Θ with the class of measurements (EVMs) on the input classical GPT $(\ell^1(\Theta), \ell^1_+(\Theta), \mathcal{P}(\Theta))$. Based on this correspondence Theorem 1 straightforwardly applies to the post-processing order of statistical experiments.

In the rest of this section, we prove Theorems 1.1, 1.2, and 2 in Sections 3.1, 3.2, and 3.3, respectively.

3.1. Proof of Theorem 1.1

The proof is split into some lemmas. We first establish some general properties of the order dimensions necessary for the proof.

Definition 3. Let $(S, \preceq_1)$ and $(T, \preceq_2)$ be posets. A map $f \colon S \to T$ is called an *order embedding* from $(S, \preceq_1)$ into $(T, \preceq_2)$ if $x \preceq_1 y$ if and only if $f(x) \preceq_2 f(y)$ for every $x, y \in S$. An order embedding is necessarily an injection. If such an order embedding

Infinite dimensionality of the post-processing order of measurements 12

exists, $(S, \preceq_1)$ is said to be *embeddable* into $(T, \preceq_2)$. If $(S, \preceq_1)$ is embeddable into $(T, \preceq_2)$, $(S, \preceq_1)$ is order isomorphic to a subset of T equipped with the restriction order of $\preceq_2$.

The following lemma is implicit in the literature [18, 19], while here we give a proof for completeness.

Lemma 1. *Let $(S, \preceq_1)$ and $(T, \preceq_2)$ be posets. Suppose that $(S, \preceq_1)$ is embeddable into $(T, \preceq_2)$. Then $\dim_{\mathrm{ord}}(S, \preceq_1) \leq \dim_{\mathrm{ord}}(T, \preceq_2)$ and $\dim_{\mathrm{ord}, \mathbb{R}}(S, \preceq_1) \leq \dim_{\mathrm{ord}, \mathbb{R}}(T, \preceq_2)$ hold.*

Proof. Let $g \colon S \to T$ be an order embedding and let $\mathcal{L}$ be a family of total orders on T such that $\preceq_2 = \bigcap \mathcal{L}$ and $|\mathcal{L}| = \dim_{\mathrm{ord}}(T, \preceq_2)$. For each $L \in \mathcal{L}$, we define a binary relation $g^{-1}(L)$ on S by

$$x g^{-1}(L) y :\overset{\mathrm{def.}}{\Leftrightarrow} g(x) L g(y) \quad (x, y \in S).$$

Then, since g is an injection, each $g^{-1}(L)$ $(L \in \mathcal{L})$ is a total order on S. Moreover for every $x, y \in S$ we have

$$\begin{aligned}
x \preceq_1 y &\iff g(x) \preceq_2 g(y) \\
&\iff g(x) L g(y) \quad (\forall L \in \mathcal{L}) \\
&\iff x g^{-1}(L) y \quad (\forall L \in \mathcal{L}),
\end{aligned}$$

which implies that the family $\{g^{-1}(L) \mid L \in \mathcal{L}\}$ realizes $\preceq_1$. Then the first claim $\dim_{\mathrm{ord}}(S, \preceq_1) \leq \dim_{\mathrm{ord}}(T, \preceq_2)$ immediately follows from the definition of the order dimension.

Let $\mathcal{F}$ be a family of order monotones characterizing $(T, \preceq_2)$ such that $|\mathcal{F}| = \dim_{\mathrm{ord}, \mathbb{R}}(T, \preceq_2)$. Then for every $x, y \in S$ we have

$$\begin{aligned}
x \preceq_1 y &\iff g(x) \preceq_2 g(y) \\
&\iff f \circ g(x) \leq f \circ g(y) \quad (\forall f \in \mathcal{F}).
\end{aligned}$$

This implies that $\{f \circ g\}_{f \in \mathcal{F}}$ characterizes $(S, \preceq_1)$. From this the second claim $\dim_{\mathrm{ord}, \mathbb{R}}(S, \preceq_1) \leq \dim_{\mathrm{ord}, \mathbb{R}}(T, \preceq_2)$ immediately follows. $\square$

Let $((S_i, \preceq_i))_{i \in I}$ be an indexed family of posets. We define a poset, called the *direct product*, by

$$\bigotimes_{i \in I} (S_i, \preceq_i) := \left(\prod_{i \in I} S_i, \preceq \right),$$

$$(x_i)_{i \in I} \preceq (y_i)_{i \in I} :\overset{\mathrm{def.}}{\Leftrightarrow} [x_i \preceq_i y_i \quad (\forall i \in I)].$$

For a poset $(S, \preceq)$ we denote by $\mathrm{dpc}(S, \preceq)$ the minimum cardinality $|I|$ of a family $((C_i, \preceq_i))_{i \in I}$ of *chains* such that $(S, \preceq)$ is embeddable into the direct product $\bigotimes_{i \in I}(C_i, \preceq_i)$. Then it is known [37] that

$$\dim_{\mathrm{ord}}(S, \preceq) = \mathrm{dpc}(S, \preceq) \tag{13}$$

holds for every poset $(S, \preceq)$. From this we can show

Infinite dimensionality of the post-processing order of measurements 13

Lemma 2. *Let* $(S, \preceq)$ *be a poset. Then* $\dim_{\mathrm{ord}}(S, \preceq) \leq \dim_{\mathrm{ord}, \mathbb{R}}(S, \preceq)$.

Proof. Let $\mathcal{F}$ be a set of order monotones on S such that $\mathcal{F}$ characterizes $\preceq$ and $|\mathcal{F}| = \dim_{\mathrm{ord}, \mathbb{R}}(S, \preceq)$. Then the map

$$S \ni x \mapsto (f(x))_{f \in \mathcal{F}} \in \mathbb{R}^{\mathcal{F}}$$

is an order embedding from $(S, \preceq)$ into the direct product $\bigotimes_{f \in \mathcal{F}}(\mathbb{R}, \leq)$, where the order $\leq$ on the reals $\mathbb{R}$ is the usual order. Since $(\mathbb{R}, \leq)$ is a chain, the claim follows from (13). $\square$

We now consider the specific base-norm Banach space $(\ell^2(\mathbb{N}_2), \ell_+^2(\mathbb{N}_2), \mathcal{P}(\mathbb{N}_2))$ and the poset $(\mathfrak{M}_{\mathrm{fin}}(\mathcal{P}(\mathbb{N}_2)), \preceq_{\mathrm{post}})$. We can and do identify $(\ell^2(\mathbb{N}_2), \ell_+^2(\mathbb{N}_2), \mathcal{P}(\mathbb{N}_2))$ and $(\ell^\infty(\mathbb{N}_2), \ell_+^\infty(\mathbb{N}_2), u_{\mathcal{P}(\mathbb{N}_2)})$ with $(\mathbb{R}^2, \mathbb{R}_+^2, \Omega_{\mathrm{cbit}})$ and $(\mathbb{R}^2, \mathbb{R}_+^2, u_{\mathrm{cbit}})$, respectively, where

$$\mathbb{R}_+ := [0, \infty),$$
$$\Omega_{\mathrm{cbit}} := \{ (p_0, p_1) \in \mathbb{R}^2 \mid p_0, p_1 \geq 0, \ p_0 + p_1 = 1 \},$$
$$u_{\mathrm{cbit}} := (1, 1).$$

Here the duality of $\mathbb{R}^2$ and $\mathbb{R}^{2*} = \mathbb{R}^2$ is given by

$$\langle (a_0, a_1), (b_0, b_1) \rangle := a_0 b_0 + a_1 b_1 \quad ((a_0, a_1), (b_0, b_1) \in \mathbb{R}^2).$$

The set Ω_{cbit} is a line segment and corresponds to the state space of a classical bit.

Lemma 3. $(\mathfrak{M}_{\mathrm{fin}}(\Omega_{\mathrm{cbit}}), \preceq_{\mathrm{post}})$ *is embeddable into* $(\mathfrak{M}_{\mathrm{fin}}(\Omega), \preceq_{\mathrm{post}})$.

Proof. From the assumption $\dim V \geq 2$, we have $\dim V^* \geq 2$. Thus there exists an element $a_0 \in V^*$ such that (a_0, u_Ω) is linearly independent. We put

$$a := \|a_0 + \|a_0\| u_\Omega\|^{-1} (a_0 + \|a_0\| u_\Omega),$$
$$a' := u_\Omega - a.$$

Then a and a' are effects and (a, a') is linearly independent. We define a linear map $\Psi \colon \mathbb{R}^2 \to V^*$ by

$$\Psi((\alpha_0, \alpha_1)) := \alpha_0 a + \alpha_1 a' \quad ((\alpha_0, \alpha_1) \in \mathbb{R}^2).$$

Then, by the linear independence of (a, a'), Ψ is injective. Moreover, Ψ is unital and positive, i.e. $\Psi(u_{\mathrm{cbit}}) = u_\Omega$ and $\Psi(\mathbb{R}_+^2) \subseteq V_+^*$ hold.

For each EVM $\mathsf{M} = (\mathsf{M}(j))_{j=0}^{m-1}$ on Ω_{cbit}, we define

$$\Psi(\mathsf{M}) := (\Psi(\mathsf{M}(j)))_{j=0}^{m-1} \in V^{*m}.$$

Since Ψ is unital and positive, $\Psi(\mathsf{M})$ is an EVM on Ω. Then for every EVMs $\mathsf{M} = (\mathsf{M}(j))_{j=0}^{m-1}$ and $\mathsf{N} = (\mathsf{N}(k))_{k=0}^{n-1}$ on Ω_{cbit} the equivalence

$$\mathsf{M} \preceq_{\mathrm{post}} \mathsf{N} \iff \Psi(\mathsf{M}) \preceq_{\mathrm{post}} \Psi(\mathsf{N}) \tag{14}$$

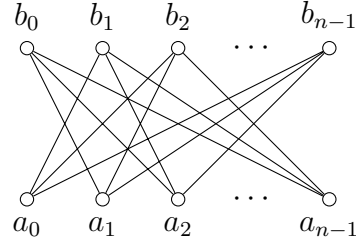*Infinite dimensionality of the post-processing order of measurements* 14

Figure 1. The Hasse diagram of the standard example $(S_n, \preceq_n)$ of an n-dimensional poset.

holds. The claim (14) can be shown as follows:

$$\mathsf{M} \preceq_{\mathrm{post}} \mathsf{N}$$

$$\Longleftrightarrow \exists (p(j|k))_{j \in \mathbb{N}_m, k \in \mathbb{N}_n} \colon \text{Markov matrix s.t.} \left[\mathsf{M}(j) = \sum_{k=0}^{n-1} p(j|k)\mathsf{N}(k) \quad (\forall j \in \mathbb{N}_m) \right]$$

$$\Longleftrightarrow \exists (p(j|k))_{j \in \mathbb{N}_m, k \in \mathbb{N}_n} \colon \text{Markov matrix s.t.} \left[\Psi(\mathsf{M}(j)) = \sum_{k=0}^{n-1} p(j|k)\Psi(\mathsf{N}(k)) \quad (\forall j \in \mathbb{N}_m) \right]$$

$$\Longleftrightarrow \Psi(\mathsf{M}) \preceq_{\mathrm{post}} \Psi(\mathsf{N}),$$

where the second equivalence follows from the injectivity of Ψ. From (14), the map

$$\mathfrak{M}_{\mathrm{fin}}(\Omega_{\mathrm{cbit}}) \ni [\mathsf{M}] \mapsto [\Psi(\mathsf{M})] \in \mathfrak{M}_{\mathrm{fin}}(\Omega)$$

is a well-defined order embedding, which completes the proof. $\qquad\square$

From Lemmas 1 and 3, the proof of the infinite dimensionality of $(\mathfrak{M}_{\mathrm{fin}}(\Omega), \preceq_{\mathrm{post}})$ reduces to the case of the classical bit space Ω_{cbit}. This is done by proving that the following *standard example of an n-dimensional poset* ([18, 19]; [20], Chapter 1, § 5) is embeddable into $(\mathfrak{M}_{\mathrm{fin}}(\Omega_{\mathrm{cbit}}), \preceq_{\mathrm{post}})$.

Definition 4 (Standard example of an n-dimensional poset)**.** For each natural number $n \geq 2$, we define a poset $(S_n, \preceq_n)$, called the standard example of an n-dimensional poset, as follows. S_n is a $2n$-element set given by $S_n := \{a_j\}_{j=0}^{n-1} \cup \{b_j\}_{j=0}^{n-1}$ and the order $\preceq_n$ is given by

$$\preceq_n := \{(a_j, a_j)\}_{j=0}^{n-1} \cup \{(b_j, b_j)\}_{j=0}^{n-1} \cup \bigcup_{j=0}^{n-1} \bigcup_{k \in \mathbb{N}_n \setminus \{j\}} \{(a_j, b_k)\}$$

(see Figure 1 for the Hasse diagram). It is known [18, 20] that $\dim_{\mathrm{ord}}(S_n, \preceq_n) = n$ holds for each $n \geq 2$.

In order to show that $(S_n, \preceq_n)$ is embeddable into $(\mathfrak{M}_{\mathrm{fin}}(\Omega_{\mathrm{cbit}}), \preceq_{\mathrm{post}})$, we need the following notion of the direct mixture of EVMs and some related lemmas.

Definition 5. Let $\mathsf{M}_j = (\mathsf{M}_j(k))_{k=0}^{m_j-1}$ $(j = 0, 1, \ldots, N-1)$ be EVMs on Ω and let $(p_j)_{j=0}^{N-1}$ be a probability distribution. Then we define a $\sum_{j=0}^{N-1} m_j =: m$-outcome EVM $\mathsf{M} = (\mathsf{M}(k))_{k=0}^{m-1}$ by

$$\mathsf{M}(k) := p_j \mathsf{M}_j \left(k - \sum_{l=0}^{j-1} m_l \right) \quad \left(\text{when } \sum_{l=0}^{j-1} m_l \le k < \sum_{l=0}^{j} m_l \right),$$

i.e.

$$\mathsf{M} := (p_0 \mathsf{M}_0(0), \ldots, p_0 \mathsf{M}_0(m_0 - 1), p_1 \mathsf{M}_1(0), \ldots, p_1 \mathsf{M}_1(m_1 - 1), \ldots,$$
$$p_{N-1} \mathsf{M}_{N-1}(0), \ldots, p_{N-1} \mathsf{M}(m_{N-1} - 1)).$$

This M is called the *direct mixture* and written as

$$\bigoplus_{j=0}^{N-1} p_j \mathsf{M}_j$$

or

$$p_0 \mathsf{M}_0 \oplus p_1 \mathsf{M}_1 \oplus \cdots \oplus p_{N-1} \mathsf{M}_{N-1}.$$

In the operational language, the measurement corresponding to the direct mixture M is realized as follows: we first generate a random number j according to the probability distribution $(p_j)_{j=0}^{N-1}$, then perform M_j which gives a measurement outcome $k \in \mathbb{N}_{m_j}$, and finally record both j and k. The reader should not confuse the direct mixture with the ordinary mixture $(\sum_{j=0}^{N-1} p_j \mathsf{M}_j(k))_{k=0}^{m'-1}$ (here we assume $m_j = m'$ for all j), which is realized when we forget j and record only k.

The state discrimination probability $P_{\mathrm{g}}(\mathcal{E}; \cdot)$ is affine with respect to the direct mixture as shown in the following lemma.

Lemma 4 (cf. [15], Proposition 14.1). *Let* $\mathsf{M}_j = (\mathsf{M}_j(k))_{k=0}^{m_j-1}$ $(j = 0, 1, \ldots, N-1)$ *be EVMs on* Ω, *let* $(p_j)_{j=0}^{N-1}$ *be a probability distribution, and let* $\mathcal{E} = (\rho_l)_{l=0}^{n-1}$ *be an ensemble on* Ω. *Then*

$$P_{\mathrm{g}} \left(\mathcal{E}; \bigoplus_{j=0}^{N-1} p_j \mathsf{M}_j \right) = \sum_{j=0}^{N-1} p_j P_{\mathrm{g}}(\mathcal{E}; \mathsf{M}_j) \tag{15}$$

holds.

Proof. By using (5), we have

$$P_{\mathrm{g}} \left(\mathcal{E}; \bigoplus_{j=0}^{N-1} p_j \mathsf{M}_j \right) = \sum_{j=0}^{N-1} \sum_{k=0}^{m_j-1} \max_{l \in \mathbb{N}_n} \langle p_j \mathsf{M}_j(k), \rho_l \rangle$$
$$= \sum_{j=0}^{N-1} p_j \sum_{k=0}^{m_j-1} \max_{l \in \mathbb{N}_n} \langle \mathsf{M}_j(k), \rho_l \rangle$$
$$= \sum_{j=0}^{N-1} p_j P_{\mathrm{g}}(\mathcal{E}; \mathsf{M}_j),$$

which proves (15). $\qquad\square$

Infinite dimensionality of the post-processing order of measurements   16

For each $(s_0, s_1) \in [0,1]^2 (= \mathcal{E}(\Omega_{\text{cbit}}))$ we write as

$$\mathsf{A}_{s_0,s_1} := ((s_0, s_1), u_{\text{cbit}} - (s_0, s_1)) = ((s_0, s_1), (1 - s_0, 1 - s_1)) \in \text{EVM}_2(\Omega_{\text{cbit}}),$$

which is the general form of a 2-outcome EVM on Ω_{cbit}.

Lemma 5. *Let $s_0, s_1, t_0, t_1 \in [0,1]$. Then $\mathsf{A}_{s_0,s_1} \preceq_{\text{post}} \mathsf{A}_{t_0,t_1}$ if and only if there exist scalars $p, q \in [0,1]$ such that $(s_0, s_1) = p(t_0, t_1) + q(1 - t_0, 1 - t_1)$, i.e. (s_0, s_1) is inside the parallelogram $(0,0)$-(t_0, t_1)-$(1,1)$-$(1 - t_0, 1 - t_1)$.*

Proof.

$\mathsf{A}_{s_0,s_1} \preceq_{\text{post}} \mathsf{A}_{t_0,t_1}$

$$\iff \exists p, p', q, q' \in [0, \infty) \quad \text{s.t.} \quad \begin{cases} p + p' = q + q' = 1, \\ (s_0, s_1) = p(t_0, t_1) + q(1 - t_0, 1 - t_1), \\ (1 - s_0, 1 - s_1) = p'(t_0, t_1) + q'(1 - t_0, 1 - t_1) \end{cases}$$

$$\iff \exists p, q \in [0,1] \quad \text{s.t.} \quad \begin{cases} (s_0, s_1) = p(t_0, t_1) + q(1 - t_0, 1 - t_1), \\ (1 - s_0, 1 - s_1) = (1 - p)(t_0, t_1) + (1 - q)(1 - t_0, 1 - t_1) \end{cases}$$

$$\iff \exists p, q \in [0,1] \quad \text{s.t.} \quad (s_0, s_1) = p(t_0, t_1) + q(1 - t_0, 1 - t_1),$$

where the final equivalence holds since the second condition

$$(1 - s_0, 1 - s_1) = (1 - p)(t_0, t_1) + (1 - q)(1 - t_0, 1 - t_1)$$

follows from the first condition

$$(s_0, s_1) = p(t_0, t_1) + q(1 - t_0, 1 - t_1). \qquad \square$$

Lemma 6. *Let $0 < s < t < 1$. Then A_{s,s^2} and A_{t,t^2} are incomparable, i.e. neither $\mathsf{A}_{s,s^2} \preceq_{\text{post}} \mathsf{A}_{t,t^2}$ nor $\mathsf{A}_{t,t^2} \preceq_{\text{post}} \mathsf{A}_{s,s^2}$ holds.*

Proof. By the strict convexity of the function $f(x) = x^2$, the points (s, s^2) and (t, t^2) are respectively below the line segment $(0,0)$-(t, t^2) and (s, s^2)-$(1,1)$. Therefore points (s, s^2) and (t, t^2) are respectively outside the parallelograms $(0,0)$-(t, t^2)-$(1,1)$-$(1 - t, 1 - t^2)$ and $(0,0)$-(s, s^2)-$(1,1)$-$(1 - s, 1 - s^2)$ (Figure 2). Hence the claim follows from Lemma 5. $\qquad \square$

We are now in a position to prove the following crucial lemma.

Lemma 7. *$(S_n, \preceq_n)$ is embeddable into $(\mathfrak{M}_{\text{fin}}(\Omega_{\text{cbit}}), \preceq_{\text{post}})$ for each natural number $n \geq 3$.*
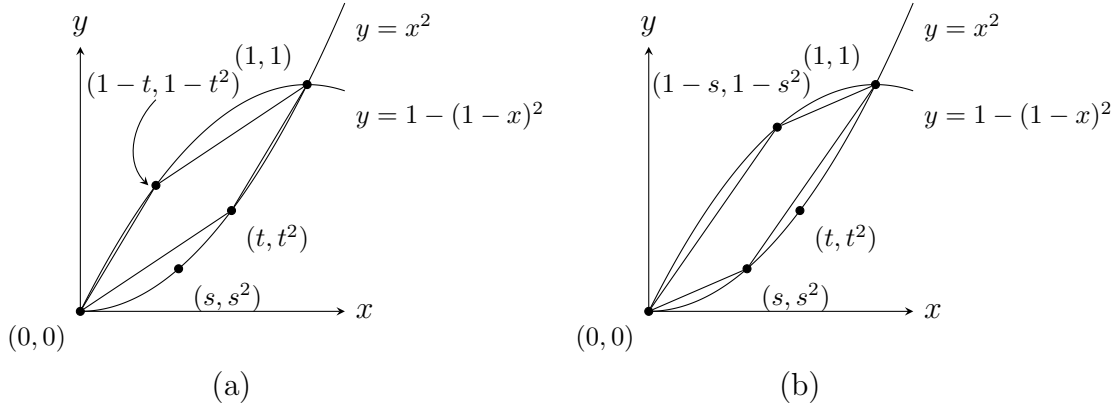
Infinite dimensionality of the post-processing order of measurements   17

(a)   (b)

Figure 2. The points (s, s^2) and (t, t^2) are respectively outside the parallelograms $(0,0)$-(t,t^2)-$(1,1)$-$(1-t, 1-t^2)$ (a) and $(0,0)$-(s,s^2)-$(1,1)$-$(1-s, 1-s^2)$ (b).

Proof. We write as $\mathsf{U} := (u_{\mathrm{cbit}})$, which is a trivial EVM on Ω_{cbit}. We define a map $f \colon S_n \to \mathfrak{M}_{\mathrm{fin}}(\Omega_{\mathrm{cbit}})$ by

$$s_j := 3^{j-n},$$

$$f(a_j) := [\mathsf{A}^{(j)}], \quad \mathsf{A}^{(j)} := \frac{1}{n}\mathsf{A}_{s_j, s_j^2} \oplus \frac{n-1}{n}\mathsf{U},$$

$$f(b_j) := [\mathsf{B}^{(j)}], \quad \mathsf{B}^{(j)} := \frac{1}{n}\mathsf{U} \oplus \bigoplus_{k \in \mathbb{N}_n \setminus \{j\}} \frac{1}{n}\mathsf{A}_{s_k, s_k^2}$$

$(0 \leq j \leq n-1)$. We establish the lemma by demonstrating that f is an order embedding. For this we have only to prove the following assertions: for every $j, k \in \mathbb{N}_n$ with $j \neq k$

 (i) $\mathsf{A}^{(j)}$ and $\mathsf{A}^{(k)}$ are incomparable;

(ii) $\mathsf{B}^{(j)}$ and $\mathsf{B}^{(k)}$ are incomparable;

(iii) $\mathsf{A}^{(j)} \preceq_{\mathrm{post}} \mathsf{B}^{(k)}$;

(iv) $\mathsf{B}^{(k)} \preceq_{\mathrm{post}} \mathsf{A}^{(j)}$ does not hold;

 (v) $\mathsf{A}^{(j)}$ and $\mathsf{B}^{(j)}$ are incomparable.

Proof of (i). From Proposition 1 and Lemma 4 we have the following equivalences:

$\mathsf{A}^{(j)} \preceq_{\mathrm{post}} \mathsf{A}^{(k)}$

$\quad \Longleftrightarrow P_{\mathrm{g}}(\mathcal{E}; \mathsf{A}^{(j)}) \leq P_{\mathrm{g}}(\mathcal{E}; \mathsf{A}^{(k)}) \quad (\forall \mathcal{E}: \text{ensemble})$

$\quad \Longleftrightarrow \frac{1}{n}P_{\mathrm{g}}(\mathcal{E}; \mathsf{A}_{s_j, s_j^2}) + \frac{n-1}{n}P_{\mathrm{g}}(\mathcal{E}; \mathsf{U}) \leq \frac{1}{n}P_{\mathrm{g}}(\mathcal{E}; \mathsf{A}_{s_k, s_k^2}) + \frac{n-1}{n}P_{\mathrm{g}}(\mathcal{E}; \mathsf{U}) \quad (\forall \mathcal{E}: \text{ensemble})$

$\quad \Longleftrightarrow P_{\mathrm{g}}(\mathcal{E}; \mathsf{A}_{s_j, s_j^2}) \leq P_{\mathrm{g}}(\mathcal{E}; \mathsf{A}_{s_k, s_k^2}) \quad (\forall \mathcal{E}: \text{ensemble})$

$\quad \Longleftrightarrow \mathsf{A}_{s_j, s_j^2} \preceq_{\mathrm{post}} \mathsf{A}_{s_k, s_k^2},$

but the last condition does not hold by Lemma 6. Hence $\mathsf{A}^{(j)} \preceq_{\mathrm{post}} \mathsf{A}^{(k)}$ does not hold. Similarly $\mathsf{A}^{(k)} \preceq_{\mathrm{post}} \mathsf{A}^{(j)}$ does not hold.

Infinite dimensionality of the post-processing order of measurements   18

Proof of (ii). By using Proposition 1 and Lemma 4 we have

$$\mathsf{B}^{(j)} \preceq_{\mathrm{post}} \mathsf{B}^{(k)}$$

$$\iff P_{\mathrm{g}}(\mathcal{E}; \mathsf{B}^{(j)}) \le P_{\mathrm{g}}(\mathcal{E}; \mathsf{B}^{(k)}) \quad (\forall \mathcal{E}: \text{ ensemble})$$

$$\iff \frac{1}{n} P_{\mathrm{g}}(\mathcal{E}; \mathsf{U}) + \sum_{l \in \mathbb{N}_n \setminus \{j\}} \frac{1}{n} P_{\mathrm{g}}(\mathcal{E}; \mathsf{A}_{s_l, s_l^2}) \le \frac{1}{n} P_{\mathrm{g}}(\mathcal{E}; \mathsf{U}) + \sum_{l \in \mathbb{N}_n \setminus \{k\}} \frac{1}{n} P_{\mathrm{g}}(\mathcal{E}; \mathsf{A}_{s_l, s_l^2}) \quad (\forall \mathcal{E}: \text{ ensemble})$$

$$\iff P_{\mathrm{g}}(\mathcal{E}; \mathsf{A}_{s_k, s_k^2}) \le P_{\mathrm{g}}(\mathcal{E}; \mathsf{A}_{s_j, s_j^2}) \quad (\forall \mathcal{E}: \text{ ensemble})$$

$$\iff \mathsf{A}_{s_k, s_k^2} \preceq_{\mathrm{post}} \mathsf{A}_{s_j, s_j^2},$$

but the last condition does not hold by Lemma 6. Hence $\mathsf{B}^{(j)} \preceq_{\mathrm{post}} \mathsf{B}^{(k)}$ does not hold. Similarly $\mathsf{B}^{(k)} \preceq_{\mathrm{post}} \mathsf{B}^{(j)}$ does not hold.

Proof of (iii). By using Proposition 1 and Lemma 4 we have

$$\mathsf{A}^{(j)} \preceq_{\mathrm{post}} \mathsf{B}^{(k)}$$

$$\iff P_{\mathrm{g}}(\mathcal{E}; \mathsf{A}^{(j)}) \le P_{\mathrm{g}}(\mathcal{E}; \mathsf{B}^{(k)}) \quad (\forall \mathcal{E}: \text{ ensemble})$$

$$\iff \frac{1}{n} P_{\mathrm{g}}(\mathcal{E}; \mathsf{A}_{s_j, s_j^2}) + \frac{n-1}{n} P_{\mathrm{g}}(\mathcal{E}; \mathsf{U}) \le \frac{1}{n} P_{\mathrm{g}}(\mathcal{E}; \mathsf{U}) + \sum_{l \in \mathbb{N}_n \setminus \{k\}} \frac{1}{n} P_{\mathrm{g}}(\mathcal{E}; \mathsf{A}_{s_l, s_l^2}) \quad (\forall \mathcal{E}: \text{ ensemble})$$

$$\iff \sum_{l \in \mathbb{N}_n \setminus \{j,k\}} (P_{\mathrm{g}}(\mathcal{E}; \mathsf{A}_{s_l, s_l^2}) - P_{\mathrm{g}}(\mathcal{E}; \mathsf{U})) \ge 0 \quad (\forall \mathcal{E}: \text{ ensemble}).$$

The last condition holds because $\mathsf{U} \preceq_{\mathrm{post}} \mathsf{M}$ for every EVM M. Thus $\mathsf{A}^{(j)} \preceq_{\mathrm{post}} \mathsf{B}^{(k)}$ holds.

Proof of (iv). By using Proposition 1 and Lemma 4 we have

$$\mathsf{B}^{(k)} \preceq_{\mathrm{post}} \mathsf{A}^{(j)}$$

$$\iff P_{\mathrm{g}}(\mathcal{E}; \mathsf{B}^{(k)}) \le P_{\mathrm{g}}(\mathcal{E}; \mathsf{A}^{(j)}) \quad (\forall \mathcal{E}: \text{ ensemble})$$

$$\iff \frac{1}{n} P_{\mathrm{g}}(\mathcal{E}; \mathsf{U}) + \sum_{l \in \mathbb{N}_n \setminus \{k\}} \frac{1}{n} P_{\mathrm{g}}(\mathcal{E}; \mathsf{A}_{s_l, s_l^2}) \le \frac{1}{n} P_{\mathrm{g}}(\mathcal{E}; \mathsf{A}_{s_j, s_j^2}) + \frac{n-1}{n} P_{\mathrm{g}}(\mathcal{E}; \mathsf{U}) \quad (\forall \mathcal{E}: \text{ ensemble})$$

$$\iff \sum_{l \in \mathbb{N}_n \setminus \{j,k\}} (P_{\mathrm{g}}(\mathcal{E}; \mathsf{A}_{s_l, s_l^2}) - P_{\mathrm{g}}(\mathcal{E}; \mathsf{U})) \le 0 \quad (\forall \mathcal{E}: \text{ ensemble}).$$

Since $P_{\mathrm{g}}(\mathcal{E}; \mathsf{A}_{s_l, s_l^2}) - P_{\mathrm{g}}(\mathcal{E}; \mathsf{U}) \ge 0$, the last condition is equivalent to

$$P_{\mathrm{g}}(\mathcal{E}; \mathsf{A}_{s_l, s_l^2}) = P_{\mathrm{g}}(\mathcal{E}; \mathsf{U}) \quad (\forall l \in \mathbb{N}_n \setminus \{j,k\}; \forall \mathcal{E}: \text{ ensemble}),$$

and therefore equivalent to $\mathsf{A}_{s_l, s_l^2} \sim_{\mathrm{post}} \mathsf{U}$ ($\forall l \in \mathbb{N}_n \setminus \{j,k\}$). The EVM $\mathsf{A}_{1,1}$ is trivial and hence post-processing equivalent to U. Since (s_l, s_l^2) is outside the line segment $(0,0)$-$(1,1)$, Lemma 5 implies that $\mathsf{A}_{s_l, s_l^2} \preceq_{\mathrm{post}} \mathsf{A}_{1,1}$, or equivalently $\mathsf{A}_{s_l, s_l^2} \preceq_{\mathrm{post}} \mathsf{U}$, does not hold. Therefore $\mathsf{B}^{(k)} \preceq_{\mathrm{post}} \mathsf{A}^{(j)}$ does not hold.

Proof of (v). We first assume $\mathsf{A}^{(j)} \preceq_{\mathrm{post}} \mathsf{B}^{(j)}$ and derive a contradiction. From the definition of the post-processing relation, the component $(s_j/n, s_j^2/n)$ of the EVM $\mathsf{A}^{(j)}$ can be written as the following conic combination of the components of $\mathsf{B}^{(j)}$:

$$\frac{1}{n}(s_j, s_j^2) = \frac{1}{n} \sum_{l \in \mathbb{N}_n \setminus \{j\}} \left[q_l(s_l, s_l^2) + r_l(1 - s_l, 1 - s_l^2) \right] + \frac{r}{n}(1,1),$$

Infinite dimensionality of the post-processing order of measurements   19

where $q_l, r_l, r \in [0, 1]$. This implies

$$(s_j, s_j^2) = \vec{u} + \vec{v}, \tag{16}$$

where

$$\vec{u} := \sum_{l: 0 \leq l \leq j-1} q_l(s_l, 0),$$

$$\vec{v} := \sum_{l: 0 \leq l \leq j-1} q_l(0, s_l^2) + \sum_{l: j+1 \leq l \leq n-1} q_l(s_l, s_l^2) + \sum_{l \in \mathbb{N}_n \setminus \{j\}} r_l(1 - s_l, 1 - s_l^2) + r(1, 1). \tag{17}$$

Since we have

$$0 \leq \sum_{l: 0 \leq l \leq j-1} q_l s_l \leq \sum_{l=0}^{j-1} s_l = \frac{3^{-n}(3^j - 1)}{2} < \frac{3^{j-n}}{2} = \frac{s_j}{2},$$

$\vec{u}$ is in the line segment $L_j := \{ (t, 0) \mid 0 \leq t \leq s_j/2 \}$. On the other hand, $\vec{v}$ is in the convex cone

$$C_j := \{ (\alpha s_{j+1}, \alpha s_{j+1}^2 + \beta) \mid \alpha, \beta \in [0, \infty) \}$$
$$= \{ (x, y) \in \mathbb{R}^2 \mid x \geq 0, y \geq s_{j+1} x \}$$

generated by (s_{j+1}, s_{j+1}^2) and $(0, 1)$, where we put $s_n := 3^{n-n} = 1$. This can be seen from that all the terms on the RHS of (17) are in C_j and C_j is closed under conic combinations. Therefore (16) implies that

$$(s_j, s_j^2) \in L_j + C_j = \{ (x + t, y) \mid x \geq 0, y \geq s_{j+1} x, 0 \leq t \leq s_j/2 \},$$

and hence there exists $0 \leq t \leq s_j/2$ such that

$$s_j^2 \geq s_{j+1}(s_j - t).$$

This implies

$$\frac{s_j}{2} \geq t \geq (1 - s_j s_{j+1}^{-1}) s_j = (1 - 3^{j-n-(j+1)+n}) s_j = \frac{2}{3} s_j > 0,$$

which is a contradiction. Thus $\mathsf{A}^{(j)} \preceq_{\text{post}} \mathsf{B}^{(j)}$ does not hold.

We next assume $\mathsf{B}^{(j)} \preceq_{\text{post}} \mathsf{A}^{(j)}$. Then from (iii) we have $\mathsf{A}^{(k)} \preceq_{\text{post}} \mathsf{B}^{(j)} \preceq_{\text{post}} \mathsf{A}^{(j)}$ and hence $\mathsf{A}^{(k)} \preceq_{\text{post}} \mathsf{A}^{(j)}$. This contradicts the incomparability of $\mathsf{A}^{(j)}$ and $\mathsf{A}^{(k)}$ which we have already proved in (i). Therefore $\mathsf{B}^{(j)} \preceq_{\text{post}} \mathsf{A}^{(j)}$ does not hold. $\square$

Proof of Theorem 1.1. From Lemma 2 we have

$$\dim_{\text{ord}}(\mathfrak{M}_{\text{fin}}(\Omega), \preceq_{\text{post}}) \leq \dim_{\text{ord}, \mathbb{R}}(\mathfrak{M}_{\text{fin}}(\Omega), \preceq_{\text{post}}). \tag{18}$$

From Lemma 1, the order embedding results in Lemmas 3 and 7 imply

$$n = \dim_{\text{ord}}(S_n, \preceq_n) \leq \dim_{\text{ord}}(\mathfrak{M}_{\text{fin}}(\Omega_{\text{cbit}}), \preceq_{\text{post}}) \leq \dim_{\text{ord}}(\mathfrak{M}_{\text{fin}}(\Omega), \preceq_{\text{post}}) \tag{19}$$

for each natural number $n \geq 3$. Then the claim follows from (18) and (19). $\square$

Infinite dimensionality of the post-processing order of measurements 　　　20

3.2. Proof of Theorem 1.2

Now we assume that V is separable in the norm topology.

Lemma 8. *There exists a countable family of order monotones that characterizes* $(\mathfrak{M}_{\mathrm{fin}}(\Omega), \preceq_{\mathrm{post}})$.

Proof. For the proof we explicitly construct a countable family of order monotones that characterizes the post-processing order. By the separability of V, the set of ensembles on Ω is also separable in the norm topology, which means that we can take a sequence $\mathcal{E}^{(i)} = (\rho_j^{(i)})_{j=0}^{N_i-1}$ $(i \in \mathbb{N})$ of ensembles on Ω such that for every ensemble $\mathcal{E} = (\rho_j)_{j=0}^{N-1}$ and every $\epsilon > 0$ there exists some $i \in \mathbb{N}$ satisfying $N_i = N$ and

$$\|\mathcal{E} - \mathcal{E}^{(i)}\| := \sum_{j=0}^{N-1} \|\rho_j - \rho_j^{(i)}\| < \epsilon.$$

We now prove that the countable family $\{P_{\mathrm{g}}(\mathcal{E}^{(i)}; \cdot)\}_{i \in \mathbb{N}}$ of order monotones characterizes $(\mathfrak{M}_{\mathrm{fin}}(\Omega), \preceq_{\mathrm{post}})$. For this, from Proposition 1, we have only to prove that for every EVMs $\mathsf{M} = (\mathsf{M}(j))_{j=0}^{m-1}$ and $\mathsf{N} = (\mathsf{N}(k))_{k=0}^{n-1}$ on Ω,

$$P_{\mathrm{g}}(\mathcal{E}^{(i)}; \mathsf{M}) \le P_{\mathrm{g}}(\mathcal{E}^{(i)}; \mathsf{N}) \quad (\forall i \in \mathbb{N}) \tag{20}$$

implies

$$P_{\mathrm{g}}(\mathcal{E}; \mathsf{M}) \le P_{\mathrm{g}}(\mathcal{E}; \mathsf{N}) \quad (\forall \mathcal{E}: \text{ ensemble}). \tag{21}$$

Assume (20). We take an arbitrary ensemble $\mathcal{E} = (\rho_j)_{j=0}^{N-1}$ on Ω and $\epsilon > 0$. By the density of $\{\mathcal{E}^{(i)}\}_{i \in \mathbb{N}}$ there exists some $i \in \mathbb{N}$ such that $N_i = N$ and $\|\mathcal{E} - \mathcal{E}^{(i)}\| < \epsilon$. Then from the definitions of $P_{\mathrm{g}}(\mathcal{E}; \mathsf{M})$ and $\preceq_{\mathrm{post}}$ we have

$$\begin{aligned}
P_{\mathrm{g}}(\mathcal{E}; \mathsf{M}) &= \sup_{\mathsf{A} \in \mathrm{EVM}(N; \mathsf{M})} \sum_{j=0}^{N-1} \langle \mathsf{A}(j), \rho_j \rangle \\
&= \sup_{\mathsf{A} \in \mathrm{EVM}(N; \mathsf{M})} \sum_{j=0}^{N-1} \left(\langle \mathsf{A}(j), \rho_j^{(i)} \rangle + \langle \mathsf{A}(j), \rho_j - \rho_j^{(i)} \rangle \right) \\
&\le \sup_{\mathsf{A} \in \mathrm{EVM}(N; \mathsf{M})} \sum_{j=0}^{N-1} \left(\langle \mathsf{A}(j), \rho_j^{(i)} \rangle + \|\rho_j - \rho_j^{(i)}\| \right) \tag{22} \\
&\le \sup_{\mathsf{A} \in \mathrm{EVM}(N; \mathsf{M})} \left(\sum_{j=1}^{N} \langle \mathsf{A}(j), \rho_j^{(i)} \rangle + \epsilon \right) \\
&= P_{\mathrm{g}}(\mathcal{E}^{(i)}; \mathsf{M}) + \epsilon, \tag{23}
\end{aligned}$$

where in deriving (22) we used the inequality

$$|\langle f, x \rangle| \le \|f\| \|x\| \quad (f \in V^*, x \in V)$$

Infinite dimensionality of the post-processing order of measurements   21

and $\|\mathsf{A}(j)\| \leq 1$. By replacing M, $\mathcal{E}$, and $\mathcal{E}^{(i)}$ in the above argument with N, $\mathcal{E}^{(i)}$, and $\mathcal{E}$, respectively, we also obtain

$$P_{\mathrm{g}}(\mathcal{E}^{(i)}; \mathsf{N}) \leq P_{\mathrm{g}}(\mathcal{E}; \mathsf{N}) + \epsilon. \tag{24}$$

From (20), (23), and (24), we have

$$P_{\mathrm{g}}(\mathcal{E}; \mathsf{M}) \leq P_{\mathrm{g}}(\mathcal{E}; \mathsf{N}) + 2\epsilon. \tag{25}$$

Since $\epsilon > 0$ is arbitrary, (25) implies $P_{\mathrm{g}}(\mathcal{E}; \mathsf{M}) \leq P_{\mathrm{g}}(\mathcal{E}; \mathsf{N})$, which completes the proof. $\quad\square$

Proof of Theorem 1.2. From (18) and Theorem 1.1, we have

$$\aleph_0 \leq \dim_{\mathrm{ord}}(\mathfrak{M}_{\mathrm{fin}}(\Omega), \preceq_{\mathrm{post}}) \leq \dim_{\mathrm{ord}, \mathbb{R}}(\mathfrak{M}_{\mathrm{fin}}(\Omega), \preceq_{\mathrm{post}}). \tag{26}$$

On the other hand, Lemma 8 implies

$$\dim_{\mathrm{ord}, \mathbb{R}}(\mathfrak{M}_{\mathrm{fin}}(\Omega), \preceq_{\mathrm{post}}) \leq \aleph_0. \tag{27}$$

Then (10) follows from (26) and (27). $\quad\square$

3.3. Proof of Theorem 2

Now we assume that $\mathcal{H}$ is a complex separable Hilbert space.

Lemma 9. $(\mathfrak{M}_{\mathrm{fin}}(\mathbf{D}(\mathcal{H})), \preceq_{\mathrm{post}})$ *is embeddable into* $(\mathfrak{C}(\mathcal{H}), \preceq_{\mathrm{CP}})$.

Proof. For each EVM $\mathsf{M} = (\mathsf{M}(j))_{j=0}^{m-1} \in \mathrm{EVM}_{\mathrm{fin}}(\mathbf{D}(\mathcal{H}))$ we define a normal channel (called the quantum-classical channel) $\Gamma^{\mathsf{M}} \colon \mathbf{B}(\mathbb{C}^m) \to \mathbf{B}(\mathcal{H})$ by

$$\Gamma^{\mathsf{M}}(a) := \sum_{j=0}^{m-1} \langle \xi_j^{(m)} | a \xi_j^{(m)} \rangle \, \mathsf{M}(j) \quad (a \in \mathbf{B}(\mathbb{C}^m)),$$

where $(\xi_j^{(m)})_{j=0}^{m-1}$ is an orthonormal basis of $\mathbb{C}^m$. It is known that

$$\mathsf{M} \preceq_{\mathrm{post}} \mathsf{N} \iff \Gamma^{\mathsf{M}} \preceq_{\mathrm{CP}} \Gamma^{\mathsf{N}} \tag{28}$$

holds for every finite-outcome EVMs M and N on $\mathbf{D}(\mathcal{H})$ ([38], Proposition 1). From (28) it readily follows that the map

$$\mathfrak{M}_{\mathrm{fin}}(\mathbf{D}(\mathcal{H})) \ni [\mathsf{M}] \mapsto [\Gamma^{\mathsf{M}}] \in \mathfrak{C}(\mathcal{H})$$

is a well-defined order embedding, which proves the claim. $\quad\square$

The proof of the following lemma is almost parallel to that of Lemma 8.

Lemma 10. *There exists a countable family of order monotones that characterizes* $(\mathfrak{C}(\mathcal{H}), \preceq_{\mathrm{CP}})$.

Infinite dimensionality of the post-processing order of measurements 22

Proof. Since $\mathcal{H} \otimes \mathbb{C}^n$ is separable for each $n \in \mathbb{N}$, we can take a dense countable family $\mathcal{E}^{(n,i)} = (\rho_k^{(n,i)})_{k=0}^{N_{n,i}-1}$ $(i \in \mathbb{N})$ of ensembles on $\mathbf{D}(\mathcal{H} \otimes \mathbb{C}^n)$ in the sense that for every ensemble $\mathcal{E} = (\rho_k)_{k=0}^{m-1}$ on $\mathbf{D}(\mathcal{H} \otimes \mathbb{C}^n)$ and every $\epsilon > 0$ there exists $i \in \mathbb{N}$ such that $N_{n,i} = m$ and

$$\|\mathcal{E} - \mathcal{E}^{(n,i)}\|_1 := \sum_{k=0}^{m-1} \|\rho_k - \rho_k^{(n,i)}\|_1 < \epsilon.$$

We establish the lemma by demonstrating that the countable family $\{P_{\mathrm{g}}^{(n)}(\mathcal{E}^{(n,i)}; \cdot)\}_{n,i \in \mathbb{N}}$ characterizes $(\mathfrak{C}(\mathcal{H}), \preceq_{\mathrm{CP}})$. From Proposition 2, it suffices to show that

$$P_{\mathrm{g}}(\mathcal{E}^{(n,i)}; \Gamma \otimes \mathrm{id}_n) \leq P_{\mathrm{g}}(\mathcal{E}^{(n,i)}; \Lambda \otimes \mathrm{id}_n) \quad (\forall i \in \mathbb{N}) \tag{29}$$

implies

$$P_{\mathrm{g}}(\mathcal{E}; \Gamma \otimes \mathrm{id}_n) \leq P_{\mathrm{g}}(\mathcal{E}; \Lambda \otimes \mathrm{id}_n) \quad (\forall \mathcal{E}: \text{ensemble on } \mathbf{D}(\mathcal{H} \otimes \mathbb{C}^n)) \tag{30}$$

for every $n \in \mathbb{N}$ and every channels $\Gamma \colon \mathbf{B}(\mathcal{K}) \to \mathbf{B}(\mathcal{H})$ and $\Lambda \colon \mathbf{B}(\mathcal{J}) \to \mathbf{B}(\mathcal{H})$.

Assume (29). Let $\mathcal{E} = (\rho_k)_{k=0}^{m-1}$ be an arbitrary ensemble on $\mathbf{D}(\mathcal{H} \otimes \mathbb{C}^n)$. For any $\epsilon > 0$ we can take $i \in \mathbb{N}$ such that $N_{n,i} = m$ and $\|\mathcal{E} - \mathcal{E}^{(n,i)}\|_1 < \epsilon$. Then we have

$$P_{\mathrm{g}}(\mathcal{E}; \Gamma \otimes \mathrm{id}_n)$$

$$= \sup_{\mathsf{M} \in \mathrm{EVM}_m(\mathbf{D}(\mathcal{K} \otimes \mathbb{C}^n))} \sum_{k=0}^{m-1} \langle (\Gamma \otimes \mathrm{id}_n)(\mathsf{M}(k)), \rho_k \rangle$$

$$= \sup_{\mathsf{M} \in \mathrm{EVM}_m(\mathbf{D}(\mathcal{K} \otimes \mathbb{C}^n))} \sum_{k=0}^{m-1} \left(\langle (\Gamma \otimes \mathrm{id}_n)(\mathsf{M}(k)), \rho_k^{(n,i)} \rangle + \langle (\Gamma \otimes \mathrm{id}_n)(\mathsf{M}(k)), \rho_k - \rho_k^{(n,i)} \rangle \right)$$

$$\leq \sup_{\mathsf{M} \in \mathrm{EVM}_m(\mathbf{D}(\mathcal{K} \otimes \mathbb{C}^n))} \sum_{k=0}^{m-1} \left(\langle (\Gamma \otimes \mathrm{id}_n)(\mathsf{M}(k)), \rho_k^{(n,i)} \rangle + \|\rho_k - \rho_k^{(n,i)}\|_1 \|(\Gamma \otimes \mathrm{id}_n)(\mathsf{M}(k))\| \right)$$

$$\leq \sup_{\mathsf{M} \in \mathrm{EVM}_m(\mathbf{D}(\mathcal{K} \otimes \mathbb{C}^n))} \left(\sum_{k=0}^{m-1} \langle (\Gamma \otimes \mathrm{id}_n)(\mathsf{M}(k)), \rho_k^{(n,i)} \rangle \right) + \epsilon \tag{31}$$

$$= P_{\mathrm{g}}(\mathcal{E}^{(n,i)}; \Gamma \otimes \mathrm{id}_n) + \epsilon, \tag{32}$$

where we used $\|(\Gamma \otimes \mathrm{id}_n)(\mathsf{M}(k))\| \leq 1$ in deriving (31). By replacing $\mathcal{E}$, $\mathcal{E}^{(n,i)}$, and Γ in the above argument with $\mathcal{E}^{(n,i)}$, $\mathcal{E}$, and Λ, respectively, we also obtain

$$P_{\mathrm{g}}(\mathcal{E}^{(n,i)}; \Lambda \otimes \mathrm{id}_n) \leq P_{\mathrm{g}}(\mathcal{E}; \Lambda \otimes \mathrm{id}_n) + \epsilon. \tag{33}$$

From (29), (32), and (33), we have

$$P_{\mathrm{g}}(\mathcal{E}; \Gamma \otimes \mathrm{id}_n) \leq P_{\mathrm{g}}(\mathcal{E}; \Lambda \otimes \mathrm{id}_n) + 2\epsilon.$$

Since $\epsilon > 0$ is arbitrary, this implies $P_{\mathrm{g}}(\mathcal{E}; \Gamma \otimes \mathrm{id}_n) \leq P_{\mathrm{g}}(\mathcal{E}; \Lambda \otimes \mathrm{id}_n)$, which completes the proof of (29) $\Longrightarrow$ (30). $\square$

Proof of Theorem 2. Since $\mathbf{T}_{\mathrm{sa}}(\mathcal{H})$ is separable in the trace norm topology and $\dim \mathbf{T}_{\mathrm{sa}}(\mathcal{H}) \geq 4$, the first claim (11) follows from Theorem 1.2.

Now we prove the second claim (12). Lemma 2 implies

$$\dim_{\mathrm{ord}}(\mathfrak{C}(\mathcal{H}), \preceq_{\mathrm{CP}}) \leq \dim_{\mathrm{ord}, \mathbb{R}}(\mathfrak{C}(\mathcal{H}), \preceq_{\mathrm{CP}}). \tag{34}$$

From Lemmas 1 and 9 and (11) we have

$$\aleph_0 = \dim_{\mathrm{ord}}(\mathfrak{M}_{\mathrm{fin}}(\mathbf{D}(\mathcal{H})), \preceq_{\mathrm{post}}) \leq \dim_{\mathrm{ord}}(\mathfrak{C}(\mathcal{H}), \preceq_{\mathrm{CP}}). \tag{35}$$

From Lemma 10 we also have

$$\dim_{\mathrm{ord}, \mathbb{R}}(\mathfrak{C}(\mathcal{H}), \preceq_{\mathrm{CP}}) \leq \aleph_0. \tag{36}$$

Then (12) follows from (34), (35), and (36). $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

4. Conclusion

In this paper we have evaluated the order and order monotone dimensions of the post-processing orders of measurements on an arbitrary non-trivial GPT Ω (Theorem 1) and of quantum channels with a fixed input Hilbert space (Theorem 2). We found that all of these order dimensions are infinite. Our results reveal that the post-processing order of measurements or quantum channels is qualitatively more complex than any order with a finite dimension, such as the adiabatic accessibility relation in thermodynamics or the LOCC convertibility relation of finite-dimensional bipartite pure states.

In the crucial step of the proof, we have explicitly constructed an order embedding from the standard example $(S_n, \preceq_n)$ of an n-dimensional poset into the poset $(\mathfrak{M}_{\mathrm{fin}}(\Omega_{\mathrm{cbit}}), \preceq_{\mathrm{post}})$ of the equivalence classes of finite-outcome EVMs on the classical bit space Ω_{cbit} for every $n \geq 3$ (Lemma 7). We also note that the BSS-type theorems (Propositions 1 and 2) played important role in the proofs of Lemmas 7, 8, and 10.

As mentioned in the introduction, we can find many other important non-total orders in physics and quantum information, especially in quantum resource theories. The present work is just the first step to evaluate the dimensions of these kinds of orders and it would be an interesting future work to investigate other orders in quantum information theory from the standpoint of the order dimension.

Acknowledgments

Appendix A. Proof of Proposition 1

In this appendix, we prove the BSS theorem for EVMs (Proposition 1).

For the proof we use the following weak$*$ topology of $(V^*)^n$.

Infinite dimensionality of the post-processing order of measurements   24

Definition 6. Let $(V, \|\cdot\|)$ be a Banach space and let $n \in \mathbb{N}$ be a natural number. The product linear space V^n equipped with the norm

$$\|(x_k)_{k=0}^{n-1}\| := \sum_{k=0}^{n-1} \|x_k\| \quad ((x_k)_{k=0}^{n-1} \in V^n)$$

is a Banach space. We can and do identify the continuous dual $(V^n)^*$ with the product linear space $(V^*)^n$ by the duality

$$\langle (f_k)_{k=0}^{n-1}, (x_k)_{k=0}^{n-1} \rangle := \sum_{k=0}^{n-1} \langle f_k, x_k \rangle \quad ((x_k)_{k=0}^{n-1} \in V^n, (f_k)_{k=0}^{n-1} \in (V^*)^n).$$

By this identification the uniform norm on $(V^*)^n$ is given by

$$\|(f_k)_{k=0}^{n-1}\| := \max_{0 \leq k \leq n-1} \|f_k\| \quad ((f_k)_{k=0}^{n-1} \in (V^*)^n).$$

The weak* topology [39, 40] $\sigma((V^*)^n, V^n)$ on $(V^*)^n$ is the weakest topology on $(V^*)^n$ such that the linear functional $(V^*)^n \ni (f_k)_{k=0}^{n-1} \mapsto \langle (f_k)_{k=0}^{n-1}, (x_k)_{k=0}^{n-1} \rangle \in \mathbb{R}$ is continuous for all $(x_k)_{k=0}^{n-1} \in V^n$. The weak* topology is a locally convex Hausdorff topology [39, 40] on $(V^*)^n$. For every weakly* continuous linear functional $\xi \colon (V^*)^n \to \mathbb{R}$ there exists a unique $(x_k)_{k=0}^{n-1} \in V^n$ such that

$$\xi((f_k)_{k=0}^{n-1}) = \langle (f_k)_{k=0}^{n-1}, (x_k)_{k=0}^{n-1} \rangle \quad ((f_k)_{k=0}^{n-1} \in (V^*)^n).$$

Proof of Proposition 1. We write as $\mathsf{M} = (\mathsf{M}(j))_{j=0}^{m-1}$ and $\mathsf{N} = (\mathsf{N}(k))_{k=0}^{n-1}$. First assume $\mathsf{M} \preceq_{\mathrm{post}} \mathsf{N}$. Then by the transitivity of $\preceq_{\mathrm{post}}$ we have $\mathrm{EVM}(N; \mathsf{M}) \subseteq \mathrm{EVM}(N; \mathsf{N})$ for every natural number N. Therefore for every ensemble $\mathcal{E} = (\rho_l)_{l=0}^{N-1}$ we have

$$P_{\mathrm{g}}(\mathcal{E}; \mathsf{M}) = \sup_{\mathsf{A} \in \mathrm{EVM}(N; \mathsf{M})} \sum_{l=0}^{N-1} \langle \mathsf{A}(l), \rho_l \rangle \leq \sup_{\mathsf{A} \in \mathrm{EVM}(N; \mathsf{N})} \sum_{l=0}^{N-1} \langle \mathsf{A}(l), \rho_l \rangle = P_{\mathrm{g}}(\mathcal{E}; \mathsf{N}),$$

which proves the "only if" part of the claim.

To show the converse implication, we assume that $\mathsf{M} \preceq_{\mathrm{post}} \mathsf{N}$ does not hold and find an ensemble $\mathcal{E}$ satisfying $P_{\mathrm{g}}(\mathcal{E}; \mathsf{M}) > P_{\mathrm{g}}(\mathcal{E}; \mathsf{N})$. From the definition of $\mathrm{EVM}(m; \mathsf{N})$ the assumption implies $\mathsf{M} \notin \mathrm{EVM}(m; \mathsf{N})$. Since $\mathrm{EVM}(m; \mathsf{N})$ is the image of the weakly* continuous affine map

$$\mathbf{Markov}(m, n) \ni (p(j|k))_{j \in \mathbb{N}_m, k \in \mathbb{N}_n} \mapsto \left(\sum_{k=0}^{n-1} p(j|k) \mathsf{N}(k) \right)_{j=0}^{m-1} \in (V^*)^m,$$

where $\mathbf{Markov}(m, n) \subseteq \mathbb{R}^{m \times n}$ is the set of $(m \times n)$-Markov matrices, and $\mathbf{Markov}(m, n)$ is a compact convex set, $\mathrm{EVM}(m; \mathsf{N})$ is a weakly* compact convex subset of $(V^*)^m$. Therefore by the Hahn-Banach separation theorem [39, 40] there exists an element $(v_j)_{j=0}^{m-1} \in V^m$ such that

$$\sum_{j=0}^{m-1} \langle \mathsf{M}(j), v_j \rangle > \sup_{\mathsf{A} \in \mathrm{EVM}(m; \mathsf{N})} \sum_{j=0}^{m-1} \langle \mathsf{A}(j), v_j \rangle. \tag{A.1}$$

Infinite dimensionality of the post-processing order of measurements   25

Since the positive cone V_+ generates V, for each $0 \leq j \leq m-1$ we may write as $v_j = v_j^+ - v_j^-$ for some $v_j^\pm \in V_+$. We define an ensemble $\mathcal{E} = (\rho_j)_{j=0}^{m-1}$ by

$$v^- := \sum_{j=0}^{m-1} v_j^-,$$

$$c := \sum_{j=0}^{m-1} \langle u_\Omega, v_j + v^- \rangle = \sum_{j=0}^{m-1} \|v_j^+\| + (m-1) \sum_{j=0}^{m-1} \|v_j^-\|,$$

$$\rho_j := c^{-1} \left(v_j + v^- \right) = c^{-1} \left(v_j^+ + \sum_{k \in \mathbb{N}_m \setminus \{j\}} v_k^- \right) \in V_+.$$

Here $c > 0$ holds since (A.1) implies that at least one $v_j^\pm$ is non-zero. Note also that $m \geq 2$ holds since $m = 1$ implies $\mathsf{M} = (u_\Omega) \preceq_{\text{post}} \mathsf{N}$. Then we have

$$P_{\text{g}}(\mathcal{E}; \mathsf{M}) \geq \sum_{j=0}^{m-1} \langle \mathsf{M}(j), \rho_j \rangle \quad (\because \mathsf{M} \in \text{EVM}(m; \mathsf{M}))$$

$$= c^{-1} \sum_{j=0}^{m-1} \langle \mathsf{M}(j), v_j \rangle + c^{-1} \sum_{j=0}^{m-1} \langle \mathsf{M}(j), v^- \rangle$$

$$= c^{-1} \sum_{j=0}^{m-1} \langle \mathsf{M}(j), v_j \rangle + c^{-1} \langle u_\Omega, v^- \rangle$$

$$> c^{-1} \sup_{\mathsf{A} \in \text{EVM}(m; \mathsf{N})} \sum_{j=0}^{m-1} \langle \mathsf{A}(j), v_j \rangle + c^{-1} \langle u_\Omega, v^- \rangle \quad (\because \text{(A.1)})$$

$$= \sup_{\mathsf{A} \in \text{EVM}(m; \mathsf{N})} \sum_{j=0}^{m-1} \langle \mathsf{A}(j), \rho_j \rangle$$

$$= P_{\text{g}}(\mathcal{E}; \mathsf{N}),$$

which completes the proof. $\square$

Remark 5. As we can see from the construction of $\mathcal{E}$ in the "if" part of the above proof, for m-outcome EVM M, $\mathsf{M} \preceq_{\text{post}} \mathsf{N}$ holds if and only if $P_{\text{g}}(\mathcal{E}; \mathsf{M}) \leq P_{\text{g}}(\mathcal{E}; \mathsf{N})$ for every m-element ensemble $\mathcal{E}$ [11].

Appendix B. Proof of Proposition 2

In this appendix we prove Proposition 2. The proof is based on the following BSS-type theorem for channels.

Lemma 11 ([35], Proposition 2). *Let* $\Gamma \colon \mathbf{B}(\mathcal{K}) \to \mathbf{B}(\mathcal{H})$ *and* $\Lambda \colon \mathbf{B}(\mathcal{J}) \to \mathbf{B}(\mathcal{H})$ *be normal unital positive maps. Then the following conditions are equivalent.*

(i) $P_{\text{g}}(\mathcal{E}; \Gamma) \leq P_{\text{g}}(\mathcal{E}; \Lambda)$ *holds for every ensemble* $\mathcal{E}$ *on* $\mathbf{D}(\mathcal{H})$.

(ii) *For every $m \in \mathbb{N}$ and every EVM $\mathsf{M} = (\mathsf{M}(j))_{j=0}^{m-1}$ on $\mathbf{D}(\mathcal{K})$, there exists an EVM $\mathsf{N} = (\mathsf{N}(j))_{j=0}^{m-1}$ on $\mathbf{D}(\mathcal{J})$ such that*

$$\Gamma(\mathsf{M}(j)) = \Lambda(\mathsf{N}(j)) \quad (j \in \mathbb{N}_m).$$

Proof of Proposition 2. Assume (i) and take a normal channel $\Psi \colon \mathbf{B}(\mathcal{K}) \to \mathbf{B}(\mathcal{J})$ such that $\Gamma = \Lambda \circ \Psi$. Then for every $n \in \mathbb{N}$ and every ensemble $\mathcal{E} = (\rho_j)_{j=0}^{m-1}$ on $\mathbf{D}(\mathcal{H} \otimes \mathbb{C}^n)$ we have

$$
\begin{aligned}
P_{\mathrm{g}}(\mathcal{E}; \Gamma \otimes \mathrm{id}_n) &= \sup_{\mathsf{M} \in \mathrm{EVM}_m(\mathbf{D}(\mathcal{K} \otimes \mathbb{C}^n))} \sum_{j=0}^{m-1} \mathrm{tr}[\rho_j (\Gamma \otimes \mathrm{id}_n)(\mathsf{M}(j))] \\
&= \sup_{\mathsf{M} \in \mathrm{EVM}_m(\mathbf{D}(\mathcal{K} \otimes \mathbb{C}^n))} \sum_{j=0}^{m-1} \mathrm{tr}[\rho_j (\Lambda \circ \Psi \otimes \mathrm{id}_n)(\mathsf{M}(j))] \\
&= \sup_{\mathsf{M} \in \mathrm{EVM}_m(\mathbf{D}(\mathcal{K} \otimes \mathbb{C}^n))} \sum_{j=0}^{m-1} \mathrm{tr}[\rho_j (\Lambda \otimes \mathrm{id}_n) \circ (\Psi \otimes \mathrm{id}_n)(\mathsf{M}(j))] \\
&\leq P_{\mathrm{g}}(\mathcal{E}; \Lambda \otimes \mathrm{id}_n),
\end{aligned}
$$

where the last inequality follows because $((\Psi \otimes \mathrm{id}_n)(\mathsf{M}(j)))_{j=0}^{m-1}$ is an EVM on $\mathbf{D}(\mathcal{J} \otimes \mathbb{C}^n)$. Thus (ii) holds.

Conversely assume (ii). We first prove $\Gamma \preceq_{\mathrm{CP}} \Lambda$ when $d := \dim \mathcal{K}$ is finite. We take orthonormal bases $(\xi_j)_{j \in \mathbb{N}_d}$ and $(e_j)_{j \in \mathbb{N}_d}$ of $\mathcal{K}$ and $\mathbb{C}^d$, respectively. We define the following maximally entangled vectors

$$\eta_{k,m} := \frac{1}{\sqrt{d}} \sum_{j \in \mathbb{N}_d} e^{\frac{2\pi i}{d} jm} \xi_j \otimes e_{j+k} \quad (k, m \in \mathbb{N}_d), \tag{B.1}$$

where the term $j + k$ means the sum modulo d. Then $(\eta_{k,m})_{k,m \in \mathbb{N}_d}$ is an orthonormal basis of $\mathcal{K} \otimes \mathbb{C}^d$ and therefore $(|\eta_{k,m}\rangle \langle \eta_{k,m}|)_{k,m \in \mathbb{N}_d}$ is an EVM on $\mathbf{D}(\mathcal{K} \otimes \mathbb{C}^d)$. From the assumption (ii) and Lemma 11, there exists an EVM $(\widetilde{\mathsf{M}}(k,m))_{k,m \in \mathbb{N}_d}$ on $\mathbf{D}(\mathcal{J} \otimes \mathbb{C}^d)$ such that

$$(\Gamma \otimes \mathrm{id}_d)(|\eta_{k,m}\rangle \langle \eta_{k,m}|) = (\Lambda \otimes \mathrm{id}_d)(\widetilde{\mathsf{M}}(k,m)) \quad (k, m \in \mathbb{N}_d). \tag{B.2}$$

From (B.1), the LHS of (B.2) is evaluated to be

$$\frac{1}{d} \sum_{j,j' \in \mathbb{N}_d} e^{\frac{2\pi i}{d}(j-j')m} \Gamma(|\xi_j\rangle \langle \xi_{j'}|) \otimes |e_{j+k}\rangle \langle e_{j'+k}|.$$

On the other hand, if we write as

$$\widetilde{\mathsf{M}}(k,m) =: \sum_{j,j' \in \mathbb{N}_d} M_{j,j'}^{(k,m)} \otimes |e_j\rangle \langle e_{j'}|,$$

then the RHS of (B.2) is given by

$$\sum_{j,j' \in \mathbb{N}_d} \Lambda(M_{j,j'}^{(k,m)}) \otimes |e_j\rangle \langle e_{j'}|.$$

Therefore (B.2) implies

$$\frac{1}{d^2}\Gamma(|\xi_j\rangle\langle\xi_{j'}|) = \frac{1}{d}e^{\frac{2\pi i}{d}(j'-j)m}\Lambda(M^{(k,m)}_{j+k,j'+k}) \quad (j,j',k,m\in\mathbb{N}_d). \tag{B.3}$$

By taking the summation over $k,m\in\mathbb{N}_d$ in (B.3), we obtain

$$\Gamma(|\xi_j\rangle\langle\xi_{j'}|) = \Lambda\circ\Phi(|\xi_j\rangle\langle\xi_{j'}|), \tag{B.4}$$

where $\Phi\colon \mathbf{B}(\mathcal{K})\to\mathbf{B}(\mathcal{J})$ is a linear map defined by

$$\Phi(|\xi_j\rangle\langle\xi_{j'}|) := \frac{1}{d}\sum_{k,m\in\mathbb{N}_d} e^{\frac{2\pi i}{d}(j'-j)m}M^{(k,m)}_{j+k,j'+k}. \tag{B.5}$$

We show that Φ is a unital CP map, from which $\Gamma = \Lambda\circ\Phi \preceq_{\mathrm{CP}} \Lambda$ follows. From the completeness condition $\sum_{k,m\in\mathbb{N}_d}\widetilde{\mathsf{M}}(k,m) = \mathbb{1}_{\mathcal{J}\otimes\mathbb{C}^d}$ we have

$$\sum_{k,m\in\mathbb{N}_d} M^{(k,m)}_{j,j'} = \delta_{j,j'}\mathbb{1}_{\mathcal{J}} \quad (j,j'\in\mathbb{N}_d), \tag{B.6}$$

where δ denotes the Kronecker delta. From (B.5) and (B.6) we obtain

$$\begin{aligned}
\Phi(\mathbb{1}_{\mathcal{K}}) &= \sum_{j\in\mathbb{N}_d}\Phi(|\xi_j\rangle\langle\xi_j|) \\
&= \frac{1}{d}\sum_{j,k,m\in\mathbb{N}_d} M^{(k,m)}_{j+k,j+k} \\
&= \frac{1}{d}\sum_{j',k,m\in\mathbb{N}_d} M^{(k,m)}_{j',j'} \\
&= \frac{1}{d}\sum_{j'\in\mathbb{N}_d}\mathbb{1}_{\mathcal{J}} \\
&= \mathbb{1}_{\mathcal{J}}.
\end{aligned}$$

Thus Φ is unital. From the positive semi-definiteness of $\widetilde{\mathsf{M}}(k,m)$, the matrix $(e^{\frac{2\pi i}{d}(j'-j)m}M^{(k,m)}_{j+k,j'+k})_{j,j'\in\mathbb{N}_d}$ is positive semi-definite and hence the Choi matrix $(\Phi(|\xi_j\rangle\langle\xi_{j'}|))_{j,j'\in\mathbb{N}_d}$ of Φ is also positive semi-definite by (B.5). Thus Φ is CP, which completes the proof of (ii) $\implies$ (i) when $\mathcal{K}$ is finite-dimensional.

Now we consider general $\mathcal{K}$. We take an orthonormal basis $(\zeta_i)_{i\in I}$ of $\mathcal{K}$ and denote by $\mathbb{F}(I)$ the family of finite subsets of the index set I, which is directed by the set inclusion relation. For each finite subset $F\in\mathbb{F}(I)$ we define projections

$$P_F := \sum_{i\in F}|\zeta_i\rangle\langle\zeta_i|, \quad P_F^{\perp} := \mathbb{1}_{\mathcal{K}} - P_F$$

and a channel

$$\Gamma_F\colon \mathbf{B}(P_F\mathcal{K})\ni a\mapsto\Gamma(a + \mathrm{tr}(\rho_F a)P_F^{\perp})\in\mathbf{B}(\mathcal{H}),$$

Infinite dimensionality of the post-processing order of measurements   28

where ρ_F is a fixed density operator on $P_F \mathcal{K}$. Since

$$\mathbf{B}(P_F \mathcal{K}) \ni a \mapsto a + \mathrm{tr}(\rho_F a) P_F^\perp \in \mathbf{B}(\mathcal{K})$$

is a channel, we have $\Gamma_F \preceq_{\mathrm{CP}} \Gamma$. Therefore, from the implication (i) $\implies$ (ii), for every $n \in \mathbb{N}$ and every ensemble $\mathcal{E}$ on $\mathbf{D}(\mathcal{H} \otimes \mathbb{C}^n)$ we obtain

$$P_{\mathrm{g}}(\mathcal{E}; \Gamma_F \otimes \mathrm{id}_n) \leq P_{\mathrm{g}}(\mathcal{E}; \Gamma \otimes \mathrm{id}_n). \tag{B.7}$$

From the assumption (ii) we also have

$$P_{\mathrm{g}}(\mathcal{E}; \Gamma \otimes \mathrm{id}_n) \leq P_{\mathrm{g}}(\mathcal{E}; \Lambda \otimes \mathrm{id}_n). \tag{B.8}$$

From (B.7) and (B.8) we obtain

$$P_{\mathrm{g}}(\mathcal{E}; \Gamma_F \otimes \mathrm{id}_n) \leq P_{\mathrm{g}}(\mathcal{E}; \Lambda \otimes \mathrm{id}_n).$$

Since $P_F \mathcal{K}$ is finite-dimensional, from what we have shown in the last paragraph, the relation $\Gamma_F \preceq_{\mathrm{CP}} \Lambda$ holds. Thus for each $F \in \mathbb{F}(I)$ there exists a channel $\Xi_F \colon \mathbf{B}(P_F \mathcal{K}) \to \mathbf{B}(\mathcal{J})$ such that $\Gamma_F = \Lambda \circ \Xi_F$. Then we have

$$\Gamma_F(P_F a P_F) = \Lambda \circ \Xi_F(P_F a P_F) \tag{B.9}$$

for each $a \in \mathbf{B}(\mathcal{H})$. Since $P_F a P_F \xrightarrow[F \in \mathbb{F}(I)]{\mathrm{uw}} a$, $P_F^\perp \xrightarrow[F \in \mathbb{F}(I)]{\mathrm{uw}} 0$, and $|\mathrm{tr}(\rho_F P_F a P_F)| \leq \|a\|$, the normality of Γ implies

$$\Gamma_F(P_F a P_F) = \Gamma(P_F a P_F) + \mathrm{tr}(\rho_F P_F a P_F)\Gamma(P_F^\perp)$$
$$\xrightarrow[F \in \mathbb{F}(I)]{\mathrm{uw}} \Gamma(a). \tag{B.10}$$

On the other hand, since $\|\Xi_F(P_F a P_F)\| \leq \|a\|$, from the ultraweak compactness of a closed ball of $\mathbf{B}(\mathcal{J})$ (Banach-Alaoglu theorem) and Tychonoff's theorem, there exists a subnet $(\Xi_{F(\alpha)})_{\alpha \in A}$ such that the ultraweak limit

$$\underset{\alpha \in A}{\mathrm{uw\text{-}lim}}\, \Xi_{F(\alpha)}(P_{F(\alpha)} a P_{F(\alpha)}) =: \Xi(a)$$

exists for every $a \in \mathbf{B}(\mathcal{K})$. Since the map

$$\mathbf{B}(\mathcal{K}) \ni a \mapsto \Xi_F(P_F a P_F) \in \mathbf{B}(\mathcal{J})$$

is a channel for each $F \in \mathbb{F}(I)$, so is $\Xi \colon \mathbf{B}(\mathcal{K}) \to \mathbf{B}(\mathcal{J})$. Thus from (B.9) and (B.10), for each $a \in \mathbf{B}(\mathcal{K})$ we obtain

$$\Gamma(a) = \underset{\alpha \in A}{\mathrm{uw\text{-}lim}}\, \Gamma_{F(\alpha)}(P_{F(\alpha)} a P_{F(\alpha)})$$
$$= \underset{\alpha \in A}{\mathrm{uw\text{-}lim}}\, \Lambda \circ \Xi_{F(\alpha)}(P_{F(\alpha)} a P_{F(\alpha)})$$
$$= \Lambda \circ \Xi(a),$$

where the last equality follows from the normality of Λ. Therefore we have $\Gamma = \Lambda \circ \Xi \preceq_{\mathrm{CP}} \Lambda$, which completes the proof of (ii) $\implies$ (i). $\square$

Infinite dimensionality of the post-processing order of measurements 29

Appendix C. Compactness principle for order dimension

In this appendix, we prove the following proposition.

Proposition 3 ([36])**.** *Let $(S, \preceq)$ be a poset and let $\mathbb{F}(S)$ denote the family of finite subsets of S. Then the following assertions hold.*

1. *For every $n \in \mathbb{N}$, the equivalence*

$$\dim_{\mathrm{ord}}(S, \preceq) \leq n \iff [\dim_{\mathrm{ord}}(F, \preceq|_F) \leq n \quad (\forall F \in \mathbb{F}(S))] \tag{C.1}$$

   holds.

2. *If $\dim_{\mathrm{ord}}(S, \preceq)$ is finite, then there exists a finite subset $F \in \mathbb{F}(S)$ such that $\dim_{\mathrm{ord}}(F, \preceq|_F) = \dim_{\mathrm{ord}}(S, \preceq)$.*

3. *$\dim_{\mathrm{ord}}(S, \preceq)$ is infinite if and only if for every $n \in \mathbb{N}$ there exists a finite subset $F \in \mathbb{F}(S)$ such that $\dim_{\mathrm{ord}}(F, \preceq|_F) > n$.*

For the proof of this proposition, we use the following product topology on the power set.

Definition 7. Let X be a set and let $\mathfrak{P}(X)$ denote the power set (i.e. the family of all the subsets of X). As usual we identify $\mathfrak{P}(X)$ with the product set $\mathbb{N}_2^X$ by the correspondence

$$\mathfrak{P}(X) \ni A \mapsto (1_A(x))_{x \in X} \in \mathbb{N}_2^X, \tag{C.2}$$

where

$$1_A(x) := \begin{cases} 1 & \text{if } x \in A; \\ 0 & \text{if } x \notin A. \end{cases}$$

From the product topology on $\mathbb{N}_2^X$ of the discrete topology on $\mathbb{N}_2$, the bijection (C.2) induces a topology on $\mathfrak{P}(X)$, which we just call the product topology on $\mathfrak{P}(X)$. By Tychonoff's theorem, the product topology on $\mathfrak{P}(X)$ is a compact Hausdorff topology. A net $(A_i)_{i \in I}$ on $\mathfrak{P}(X)$ converges to $A \in \mathfrak{P}(X)$ in the product topology if and only if $1_{A_i}(x) \to 1_A(x)$ for all $x \in X$, or equivalently, if and only if

$$\limsup_{i \in I} A_i = \liminf_{i \in I} A_i = A,$$

where

$$\limsup_{i \in I} A_i := \bigcap_{i \in I} \bigcup_{j \in I: j \geq i} A_j, \quad \liminf_{i \in I} A_i := \bigcup_{i \in I} \bigcap_{j \in I: j \geq i} A_j.$$

From this we can see that a net $(A_i)_{i \in I}$ on $\mathfrak{P}(X)$ is convergent if and only if $\limsup_{i \in I} A_i = \liminf_{i \in I} A_i$.

Proof of Proposition 3.

1. The "$\implies$" part of the claim (C.1) is obvious from Lemma 1. To prove the converse implication, take arbitrary $n \in \mathbb{N}$ and assume that $\dim_{\mathrm{ord}}(F, \preceq|_F) \leq n$ for all $F \in \mathbb{F}(S)$. Then for every $F \in \mathbb{F}(S)$ there exists a sequence $(L_F^{(k)})_{k=0}^{n-1}$ of total

Infinite dimensionality of the post-processing order of measurements 30

orders on F such that $\preceq|_F$ is realized by $\{L_F^{(k)}\}_{k=0}^{n-1}$. (The sequence $(L_F^{(k)})_{k=0}^{n-1}$ can contain identical elements if $\dim_{\mathrm{ord}}(F, \preceq|_F) < n$.) Since $L_F^{(k)} \subseteq F \times F \subseteq S \times S$, we may regard $(L_F^{(k)})_{F \in \mathbb{F}(S)}$ as a net on $\mathfrak{P}(S \times S)$. Then by the compactness of the product topology on $\mathfrak{P}(S \times S)$, there exist subnets $(L_{F(i)}^{(k)})_{i \in I}$ $(k \in \mathbb{N}_n)$ such that

$$\limsup_{i \in I} L_{F(i)}^{(k)} = \liminf_{i \in I} L_{F(i)}^{(k)} =: L^{(k)} \in \mathfrak{P}(S \times S) \quad (k \in \mathbb{N}_n).$$

We now show that $\{L^{(k)}\}_{k=0}^{n-1}$ is a family of total orders on S that realizes $\preceq$, which proves the " $\Longleftarrow$ " part of the claim (C.1).

We first prove that each $L^{(k)}$ is a total order on S. Let $x, y, z \in S$ be arbitrary elements.

(Reflexivity). If $\{x\} \subseteq F$, we have $(x, x) \in L_F^{(k)}$ and therefore $(x, x) \in L_{F(i)}^{(k)}$ eventually holds. Thus $(x, x) \in L^{(k)}$.

(Antisymmetry). Assume $(x, y), (y, x) \in L^{(k)}$. Then $(x, y), (y, x) \in L_{F(i)}^{(k)}$ for some $i \in I$. By the antisymmetry of $L_{F(i)}^{(k)}$, this implies $x = y$.

(Transitivity). By the transitivity of $L_F^{(k)}$ we have

$$1_{L_F^{(k)}}(x, y) 1_{L_F^{(k)}}(y, z) \leq 1_{L_F^{(k)}}(x, z)$$

when $\{x, y, z\} \subseteq F$. Since $\{x, y, z\} \subseteq F(i)$ eventually holds for $i \in I$, we have

$$\begin{aligned}
1_{L^{(k)}}(x, y) 1_{L^{(k)}}(y, z) &= \lim_{i \in I} 1_{L_{F(i)}^{(k)}}(x, y) 1_{L_{F(i)}^{(k)}}(y, z) \\
&\leq \lim_{i \in I} 1_{L_{F(i)}^{(k)}}(x, z) \\
&= 1_{L^{(k)}}(x, z),
\end{aligned}$$

which implies the transitivity of $L^{(k)}$.

(Totality). From the totality of $L_F^{(k)}$ we have

$$1_{L_F^{(k)}}(x, y) + 1_{L_F^{(k)}}(y, x) \geq 1$$

when $\{x, y\} \subseteq F$. Therefore we have

$$1_{L^{(k)}}(x, y) + 1_{L^{(k)}}(y, x) = \lim_{i \in I}[1_{L_{F(i)}^{(k)}}(x, y) + 1_{L_{F(i)}^{(k)}}(y, x)] \geq 1,$$

which implies the totality of $L^{(k)}$.

Thus we have shown that $L^{(k)}$ is a total order on S.

We now show that $\{L^{(k)}\}_{k=0}^{n-1}$ realizes $\preceq$. For this, we have only to prove

$$x \preceq y \iff [(x, y) \in L^{(k)} \quad (\forall k \in \mathbb{N}_n)] \tag{C.3}$$

for arbitrary $x, y \in S$. Since $\{L_F^{(k)}\}_{k=0}^{n-1}$ realizes $\preceq|_F$, we have

$$1_{\preceq}(x, y) = 1_{\preceq|_F}(x, y) = \prod_{k=0}^{n-1} 1_{L_F^{(k)}}(x, y)$$

Infinite dimensionality of the post-processing order of measurements   31

when $\{x, y\} \subseteq F$. Therefore

$$\prod_{k=0}^{n-1} 1_{L^{(k)}}(x, y) = \lim_{i \in I} \prod_{k=0}^{n-1} 1_{L_{F(i)}^{(k)}}(x, y) = 1_{\preceq}(x, y),$$

which implies (C.3). Thus $\{L^{(k)}\}_{k=0}^{n-1}$ realizes $\preceq$.

2. From Lemma 1, we have $\dim_{\mathrm{ord}}(F, \preceq|_F) \leq \dim_{\mathrm{ord}}(S, \preceq) =: d \in \mathbb{N}$ for every $F \in \mathbb{F}(S)$. If there exists no $F \in \mathbb{F}(S)$ such that $\dim_{\mathrm{ord}}(F, \preceq|_F) = d$, then $\dim_{\mathrm{ord}}(F, \preceq|_F) \leq d-1$ for all $F \in \mathbb{F}(I)$ and hence (C.1) implies $d = \dim_{\mathrm{ord}}(S, \preceq) \leq d - 1$, which is a contradiction. Thus there exists $F \in \mathbb{F}(S)$ such that $\dim_{\mathrm{ord}}(F, \preceq|_F) = d$.

3. The claim is proved as

$$\dim_{\mathrm{ord}}(S, \preceq) \text{ is infinite} \iff \forall n \in \mathbb{N}, \, [\dim_{\mathrm{ord}}(S, \preceq) > n]$$
$$\iff \forall n \in \mathbb{N}, \, \exists F \in \mathbb{F}(S) \, [\dim_{\mathrm{ord}}(F, \preceq|_F) > n],$$

where the last equivalence follows from (C.1). □

References

[1] Giles R 1964 *Mathematical Foundations of Thermodynamics* (Pergamon Press)
[2] Lieb E H and Yngvason J 1999 *Phys. Rep.* **310** 1 – 96 URL https://doi.org/10.1016/S0370-1573(98)00082-9
[3] Horodecki R, Horodecki P, Horodecki M and Horodecki K 2009 *Rev. Mod. Phys.* **81**(2) 865–942 URL https://link.aps.org/doi/10.1103/RevModPhys.81.865
[4] Nielsen M A and Chuang I L 2010 *Quantum Computation and Quantum Information: 10th Anniversary Edition* (Cambridge University Press)
[5] Martens H and de Muynck W M 1990 *Found. Phys.* **20** 255–281 URL https://doi.org/10.1007/BF00731693
[6] Dorofeev S and de Graaf J 1997 *Indag. Math. (N.S.)* **8** 349 – 369 URL http://dx.doi.org/10.1016/S0019-3577(97)81815-0
[7] Buscemi F, Keyl M, D'Ariano G M, Perinotti P and Werner R F 2005 *J. Math. Phys.* **46** 082109 URL https://doi.org/10.1063/1.2008996
[8] Jenčová A, Pulmannová S and Vinceková E 2008 *Int. J. Theor. Phys.* **47** 125–148 URL http://dx.doi.org/10.1007/s10773-007-9396-0
[9] Kuramochi Y 2015 *J. Math. Phys.* **56** 102205 URL http://scitation.aip.org/content/aip/journal/jmp/56/10/10.1063/1.4934235
[10] Kuramochi Y 2016 *J. Math. Phys.* **57** 089901 URL http://scitation.aip.org/content/aip/journal/jmp/57/8/10.1063/1.4961516
[11] Guff T, McMahon N A, Sanders Y R and Gilchrist A 2021 *J. Phys. A: Math. Theor.* **54** 225301 URL https://doi.org/10.1088/1751-8121/abed67
[12] Chitambar E and Gour G 2019 *Rev. Mod. Phys.* **91**(2) 025001 URL https://link.aps.org/doi/10.1103/RevModPhys.91.025001
[13] Buscemi F 2016 *Probl. Inf. Transm.* **52** 201–213 URL https://doi.org/10.1134/S0032946016030017
[14] Skrzypczyk P and Linden N 2019 *Phys. Rev. Lett.* **122**(14) 140403 URL https://link.aps.org/doi/10.1103/PhysRevLett.122.140403
[15] Kuramochi Y 2020 Compact convex structure of measurements and its applications to simulability, incompatibility, and convex resource theory of continuous-outcome measurements (*Preprint* https://arxiv.org/abs/2002.03504v2)

Infinite dimensionality of the post-processing order of measurements   32

[16] Janotta P and Hinrichsen H 2014 *J. Phys. A: Math. Theor.* **47** 323001 URL http://stacks.iop.org/1751-8121/47/i=32/a=323001

[17] Lami L 2018 *Non-classical correlations in quantum mechanics and beyond* Ph.D. thesis URL https://ddd.uab.cat/record/187745

[18] Dushnik B and Miller E W 1941 *Am. J. Math.* **63** 600–610 URL http://www.jstor.org/stable/2371374

[19] Hiraguti T 1955 *Sci. Rep. Kanazawa Univ.* **4** 1–20 URL http://scirep.w3.kanazawa-u.ac.jp/articles/04-01-001.pdf

[20] Trotter W T 1992 *Combinatorics and partially ordered sets : dimension theory* Johns Hopkins series in the mathematical sciences (Johns Hopkins University Press)

[21] Halmos P R 1960 *Naive set theory* (Van Nostrand Company)

[22] Coecke B, Fritz T and Spekkens R W 2016 *Inf. Comput.* **250** 59–86 URL https://www.sciencedirect.com/science/article/pii/S0890540116000353

[23] Szpilrajn E 1930 *Fundam. Math.* **16** 386–389 URL https://doi.org/10.4064/fm-16-1-386-389

[24] Janotta P and Lal R 2013 *Phys. Rev. A* **87**(5) 052131 URL https://link.aps.org/doi/10.1103/PhysRevA.87.052131

[25] Le Cam L 1986 *Asymptotic Methods in Statistical Decision Theory* (Springer)

[26] Torgersen E 1991 *Comparison of Statistical Experiments* (Cambridge University Press)

[27] Stinespring W F 1955 *Proc. Amer. Math. Soc.* **6** 211–216 URL http://www.jstor.org/stable/2032342

[28] Paulsen V 2003 *Completely Bounded Maps and Operator Algebras* Cambridge Studies in Advanced Mathematics (Cambridge University Press)

[29] Takesaki M 1979 *Theory of Operator Algebras I* (Springer)

[30] Guţă M and Jenčová A 2007 *Commun. Math. Phys.* **276** 341–379 ISSN 1432-0916 URL http://dx.doi.org/10.1007/s00220-007-0340-1

[31] Kuramochi Y 2018 *J. Math. Phys.* **59** 042203 URL https://doi.org/10.1063/1.5008300

[32] Kuramochi Y 2020 *Open Systems & Information Dynamics* **27** 2050003 URL https://doi.org/10.1142/S1230161220500031

[33] Chefles A 2009 The quantum Blackwell theorem and minimum error state discrimination (*Preprint* https://arxiv.org/abs/0907.0866)

[34] Shmaya E 2005 *J. Phys. A: Math. Gen.* **38** 9717–9727 URL https://doi.org/10.1088/0305-4470/38/44/008

[35] Łuczak A 2019 *J. Math. Phys.* **60** 022203 URL https://doi.org/10.1063/1.5074187

[36] Harzheim E 1970 *Math. Nachr.* **46** 183–188 URL https://onlinelibrary.wiley.com/doi/abs/10.1002/mana.19700460115

[37] Milner E and Pouzet M 1990 *Order* **7** 101–102 URL https://doi.org/10.1007/BF00383178

[38] Heinosaari T and Miyadera T 2017 *J. Phys. A: Math. Theor.* **50** 135302 URL http://stacks.iop.org/1751-8121/50/i=13/a=135302

[39] Dunford N and Schwartz J T 1958 *Linear Operators, Part I* (Interscience, New York)

[40] Schaefer H H 1999 *Topological Vector Spaces (2nd ed.)* (Springer)