

Optimal quantum observables

Erkka Haapasalo^{1, a)} and Juha-Pekka Pellonpää^{2, b)}

¹⁾*Department of Nuclear Engineering, Kyoto University*

²⁾*Turku Centre for Quantum Physics, Department of Physics and Astronomy,
University of Turku*

(Dated: 19 July 2017)

Various forms of optimality for quantum observables described as normalized positive operator valued measures (POVMs) are studied in this paper. We give characterizations for observables that determine the values of the measured quantity with probabilistic certainty or a state of the system before or after the measurement. We investigate observables which are free from noise caused by classical post-processing, mixing, or pre-processing of quantum nature. Especially, a complete characterization of pre-processing and post-processing clean observables is given, and necessary and sufficient conditions are imposed on informationally complete POVMs within the set of pure states. We also discuss joint and sequential measurements of optimal quantum observables.

PACS numbers: 03.65.Ta, 03.67.-a

^{a)}Electronic mail: haapasalo.erkka.45e@st.kyoto-u.ac.jp

^{b)}Electronic mail: juhpollo@utu.fi

I. INTRODUCTION

A normalized positive operator valued measure (POVM) describes the statistics of the outcomes of a quantum measurement and thus we call them as observables of a quantum system. However, some observables can be considered better than the others according to different criteria: The observable may be powerful enough to differentiate between any given initial states of the system or it may be decisive enough to completely determine the state after the measurement no matter how we measure this observable. The observable may also be free from different types of noise either of classical or quantum nature or its measurement cannot be reduced to a measurement of a more informational observable from the measurement of which it can be obtained by modifying either the initial state or the outcome statistics.

We study these various notions of optimality for quantum observables and investigate how they are interrelated. An extensive review of optimal observables and new results especially dealing with post- and pre-processing are given. In this introduction, we approach these problems within a simple setting only considering discrete observables of finite-dimensional quantum systems, formally define the optimality properties outlined above, and characterize observables associated with these properties. In the rest of this paper, we give definitions of optimality in the general case involving also ‘continuous’ observables of infinite-dimensional systems and characterize the optimal observables. However, one can obtain valuable insight in this general case first by looking at the mathematically simpler discussion as follows.

As advertised, let us first consider a POVM \mathbf{M} with finitely many values (or outcomes) $\Omega = \{x_1, x_2, \dots, x_N\}$ on a finite-dimensional quantum system with the associated Hilbert space \mathcal{H} (denote $d = \dim \mathcal{H} < \infty$). This means that we do not need to go into measure theoretical or functional analytical details in this introduction. The POVM \mathbf{M} can be viewed as a collection $(\mathbf{M}_1, \mathbf{M}_2, \dots, \mathbf{M}_N)$ of positive semidefinite $d \times d$ -matrices \mathbf{M}_i such that $\sum_{i=1}^N \mathbf{M}_i$ is the identity matrix when (by fixing an orthonormal basis) we identify \mathcal{H} with \mathbb{C}^d and the bounded operators on \mathcal{H} with elements of the matrix algebra $\mathcal{M}_d(\mathbb{C})$. A state of the system is represented as a density matrix ρ , that is, a positive semidefinite matrix of trace 1, and the number $p_i = \text{tr}[\rho \mathbf{M}_i] \in [0, 1]$ is interpreted as the probability of getting an outcome x_i when a measurement of \mathbf{M} is performed and the system is in the (initial or input) state ρ . Actually, \mathbf{M} is a map which assigns to each subset X of Ω a positive matrix $\mathbf{M}(X) = \sum_{x_i \in X} \mathbf{M}_i$ so

that $\text{tr}[\rho \mathbf{M}(X)]$ is the probability of getting an outcome belonging to the set X . Especially, $\mathbf{M}(\{x_i\}) = \mathbf{M}_i$.

We fix a POVM \mathbf{M} as above and study its different optimality criteria (in the categories of discrete POVMs in finite dimensions). For that we will need another discrete POVM \mathbf{M}' , or $(\mathbf{M}'_1, \mathbf{M}'_2, \dots, \mathbf{M}'_{N'})$, which acts in a d' -dimensional Hilbert space $\mathcal{H}' \cong \mathbb{C}^{d'}$. Without restricting generality, we will assume that the matrices \mathbf{M}_i and \mathbf{M}'_j are nonzero. Namely if, for instance, $\mathbf{M}_i = 0$ then $p_i = 0$ regardless of the state ρ so the outcome x_i is never obtained and we may replace Ω by $\Omega \setminus \{x_i\}$ and similarly remove all outcomes related to zero matrices.

Write $\mathbf{M}_i = \sum_{k=1}^{m_i} \lambda_{ik} |\varphi_{ik}\rangle\langle\varphi_{ik}| = \sum_{k=1}^{m_i} |d_{ik}\rangle\langle d_{ik}|$ where the eigenvectors φ_{ik} , $k = 1, \dots, m_i$, form an orthonormal set, the eigenvalues λ_{ik} are positive (and bounded by 1), and $d_{ik} = \sqrt{\lambda_{ik}}\varphi_{ik}$. We say that m_i is the multiplicity of the outcome x_i or the rank of \mathbf{M}_i , and \mathbf{M} is of rank 1 if $m_i = 1$ for all $i = 1, \dots, N$. Recall that rank-1 observables have many important properties^{3,21,32-35}. For example, their measurements break entanglement completely between the system and its environment³³. One can define a (maximal) rank-1 refinement POVM \mathbf{M}^1 of \mathbf{M} via $\mathbf{M}_{ik}^1 = |d_{ik}\rangle\langle d_{ik}|$, $i = 1, \dots, N$, $k = 1, \dots, m_i$. Now \mathbf{M}^1 has $N^1 = \sum_{i=1}^N m_i$ outcomes and $p_i = \text{tr}[\rho \mathbf{M}_i] = \sum_{k=1}^{m_i} \text{tr}[\rho \mathbf{M}_{ik}^1]$ (i.e. \mathbf{M} is a relabeling of \mathbf{M}^1) thus showing that any measurement of \mathbf{M}^1 can be viewed as a measurement of \mathbf{M} , the so-called complete measurement, since the value space of \mathbf{M}^1 ‘contains’ also the multiplicities $k \leq m_i$ of the measurement outcomes x_i of \mathbf{M} , see³³⁻³⁵ for further properties of complete measurements.

Let then \mathcal{H}_\oplus be a Hilbert space spanned by an orthonormal basis e_{ik} where $i = 1, \dots, N$ and $k = 1, \dots, m_i$. Obviously, $\dim \mathcal{H}_\oplus = N^1$. Define a discrete normalized projection valued measure (PVM) $\mathbf{P} = (\mathbf{P}_1, \dots, \mathbf{P}_N)$ of \mathcal{H}_\oplus via $\mathbf{P}_i = \sum_{k=1}^{m_i} |e_{ik}\rangle\langle e_{ik}|$ so that $\mathbf{P}_i \mathcal{H}_\oplus$ is spanned by the vectors e_{ik} , $k = 1, \dots, m_i$, and we may write (the direct sum) $\mathcal{H}_\oplus = \bigoplus_{i=1}^N (\mathbf{P}_i \mathcal{H}_\oplus)$. Define an isometry $J : \mathcal{H} \rightarrow \mathcal{H}_\oplus$, $J = \sum_{i=1}^N \sum_{k=1}^{m_i} |e_{ik}\rangle\langle d_{ik}|$ for which $J^* \mathbf{P}_i J = \mathbf{M}_i$. Hence, $(\mathcal{H}_\oplus, J, \mathbf{P})$ is a Naimark dilation of \mathbf{M} . The dilation is minimal, i.e., the span of vectors $\mathbf{P}_i J \phi$, $i = 1, \dots, N$, $\phi \in \mathcal{H}$, is the whole \mathcal{H}_\oplus . Indeed, this follows immediately from equation $\psi = \sum_{i=1}^N \sum_{k=1}^{m_i} \langle e_{ik} | \psi \rangle e_{ik} = \sum_{i=1}^N \sum_{k=1}^{m_i} \langle e_{ik} | \psi \rangle \lambda_{ik}^{-1} \mathbf{P}_i J d_{ik}$ where $\psi \in \mathcal{H}_\oplus$. Note that one can identify \mathcal{H} with a (closed) subspace $J\mathcal{H}$ of \mathcal{H}_\oplus , equipped with the projection JJ^* from \mathcal{H}_\oplus onto $J\mathcal{H}$, and we may briefly write $\mathcal{H}_\oplus = \mathcal{H} \oplus \mathcal{H}^\perp$. Especially, any state ρ of \mathcal{H} can be viewed as a state $J\rho J^*$ of the bigger space \mathcal{H}_\oplus . By using this interpretation, any measurement of \mathbf{P}

in the subsystem's state can be viewed as a measurement of \mathbf{M} via $p_i = \text{tr}[\rho \mathbf{M}_i] = \text{tr}[J\rho J^* \mathbf{P}_i]$. Finally, we note that \mathbf{M} is a PVM if and only if J is unitary (i.e. $\{d_{ik}\}_{i,k}$ is an orthonormal basis of \mathcal{H}). In this case one can identify \mathcal{H}_\oplus with \mathcal{H} and \mathbf{P} with \mathbf{M} e.g. by setting $e_{ik} = d_{ik}$.

Remark 1. Let $(\mathcal{H}_\oplus, J, \mathbf{P})$ be a (minimal) Naimark dilation of the POVM \mathbf{M} as above. Without restricting generality, one can pick any orthonormal basis $\{e_n\}_{n=1}^\infty$ of an infinite-dimensional Hilbert space \mathcal{H}_∞ and choose $e_{ik} = e_i \otimes e_k \in \mathcal{H}_\infty \otimes \mathcal{H}_\infty$ so that \mathcal{H}_\oplus becomes a (closed) subspace of $\mathcal{H}_\infty \otimes \mathcal{H}_\infty$ and $\mathbf{P}_i = |e_i\rangle\langle e_i| \otimes \sum_{k=1}^{m_i} |e_k\rangle\langle e_k| \leq \mathbf{P}'_i \otimes I_M$ where $\mathbf{P}'_i = |e_i\rangle\langle e_i|$, $i = 1, \dots, N$, constitutes a rank-1 PVM \mathbf{P}' in an N -dimensional space $\mathcal{H}_N = \text{lin}\{|e_i\rangle | i = 1, \dots, N\}$ and $I_M = \sum_{k=1}^M |e_k\rangle\langle e_k|$ is the identity operator of $\mathcal{H}_M = \text{lin}\{|e_i\rangle | i = 1, \dots, M\}$ where $M = \max_{i \leq N} \{m_i\}$. In addition, J can be interpreted as an isometry from \mathcal{H} into $\mathcal{H}_N \otimes \mathcal{H}_M$ by the same formula $J = \sum_{i=1}^N \sum_{k=1}^{m_i} |e_i \otimes e_k\rangle\langle d_{ik}|$ and we have $\mathbf{M}_i = J^*(\mathbf{P}'_i \otimes I_M)J = \Phi_J(\mathbf{P}'_i)$ where Φ_J is a (completely positive) Heisenberg channel, $\Phi_J(B) = J^*(B \otimes I_M)J = \sum_{s=1}^M \mathbf{A}_s^* B \mathbf{A}_s$ where B is an $N \times N$ -matrix (i.e. $B \in \mathcal{M}_N(\mathbb{C})$). Hence, $(\mathcal{H}_N \otimes \mathcal{H}_M, J, \mathbf{P}' \otimes I_M)$ is a Naimark dilation of \mathbf{M} , which is minimal if and only if $m_i = M$ for all $i = 1, \dots, N$. Note that the Kraus operators $\mathbf{A}_s = \sum_{i=1}^N |e_i\rangle\langle d_{is}|$ are linearly independent, i.e. the Kraus decomposition of Φ_J is minimal. In addition, the corresponding Schrödinger channel $(\Phi_J)_*$ transforms a $d \times d$ -state ρ to the $N \times N$ -state $\rho' = \sum_{s=1}^M \mathbf{A}_s \rho \mathbf{A}_s^*$ and $p_i = \text{tr}[\rho \mathbf{M}_i] = \text{tr}[\rho' \mathbf{P}'_i] = \langle e_i | \rho' | e_i \rangle$ holds.

Davies and Lewis¹¹ introduced the concept of instrument which turned out to be crucial in developing quantum measurement theory since, besides measurement statistics, it also describes the conditional state changes due to a quantum measuring process. For example, if the measurement outcome set is finite, $\Omega = \{x_1, \dots, x_N\}$, then any (Schrödinger) instrument \mathcal{I} describing a measurement of \mathbf{M} (with the outcomes Ω), can be viewed as a collection $(\mathcal{I}_i)_{i=1}^N$ of completely positive operations on $\mathcal{M}_d(\mathbb{C})$: for any i , $\mathcal{I}_i(\rho) = \sum_s \mathbf{A}_{is} \rho \mathbf{A}_{is}^*$ (a Kraus decomposition) and the dual (Heisenberg) operation is $\mathcal{J}_i(B) = \mathcal{I}_i^*(B) = \sum_s \mathbf{A}_{is}^* B \mathbf{A}_{is}$ where B is any $d \times d$ -matrix \mathcal{I}_i . Moreover $\sum_{i=1}^N \text{tr}[\mathcal{I}_i(\rho)] = 1$ for any state ρ . Now \mathcal{I} transforms an input state ρ to a (nonnormalized) output state $\mathcal{I}_i(\rho)$ if x_i is obtained. In addition, \mathcal{I} defines the measurement outcome probabilities p_i and the corresponding POVM \mathbf{M} via $p_i = \text{tr}[\mathcal{I}_i(\rho)] = \text{tr}[\rho \mathbf{M}_i]$, i.e., using the Kraus decompositions of the operations, $\mathbf{M}_i = \sum_s \mathbf{A}_{is}^* \mathbf{A}_{is}$. Note that $\rho \mapsto \sum_{i=1}^N \mathcal{I}_i(\rho)$ is a (Schrödinger) channel which transforms any state of the system to another state of the same system. More generally, a quantum channel is

a completely positive trace-preserving (cptp) linear map between state spaces associated to quantum systems (with possibly different Hilbert spaces \mathcal{H} , \mathcal{H}') so that channels transmit quantum information between different systems. Similarly, with possibly different input and output spaces $\mathcal{H} \cong \mathbb{C}^d$ and $\mathcal{H}' \cong \mathbb{C}^{d'}$, one may also assume an initial state of \mathcal{H} to transform into conditional states of \mathcal{H}' as a result of the measurement prompting to describe the measurement through an instrument \mathcal{I} with Schrödinger operations $\mathcal{I}_i : \mathcal{M}_d(\mathbb{C}) \rightarrow \mathcal{M}_{d'}(\mathbb{C})$.

Now we are ready to introduce the following six optimality criteria for \mathbf{M} :

- (1a) \mathbf{M} *determines the future of the system (completely)* if each instrument \mathcal{I} implementing \mathbf{M} is nuclear (or preparatory), i.e., of the form $\mathcal{I}_i(\rho) = p_i \sigma_i$ where σ_i 's are density matrices (of any fixed output Hilbert space \mathcal{H}') which do not depend on the input state ρ . If the outcome x_i is obtained with the nonzero probability $p_i = \text{tr}[\rho \mathbf{M}_i]$ then the output system is in the ρ -independent state σ_i after the measurement. It can be shown that \mathbf{M} *determines the future if and only if \mathbf{M} is of rank 1*, i.e. each \mathbf{M}_i is of the form $|d_i\rangle\langle d_i|$ where $d_i \in \mathcal{H}^{21,32}$.
- (1b) \mathbf{M} *is post-processing maximal (post-processing clean)* if the condition $\mathbf{M}_i = \sum_{j=1}^{N'} p'_{ji} \mathbf{M}'_j$ for all i (where (p'_{ji}) is $N' \times N$ -probability matrix and $\mathbf{M}' = (\mathbf{M}'_1, \dots, \mathbf{M}'_{N'})$ is a POVM of the same Hilbert space $\mathcal{H}' = \mathcal{H}$) implies that $\mathbf{M}'_j = \sum_{i=1}^N p_{ij} \mathbf{M}_i$ for all j where (p_{ij}) is $N \times N'$ -probability matrix. Recall that (p_{ij}) is a probability (or stochastic or Markov) matrix if $p_{ij} \geq 0$ and $\sum_j p_{ij} = 1$ for all i . The numbers p_{ij} are transition probabilities and \mathbf{M}' is said to be a smearing of \mathbf{M} if $\mathbf{M}'_j = \sum_i p_{ij} \mathbf{M}_i$ holds. In this case, \mathbf{M} and \mathbf{M}' are jointly measurable, a joint observable being $\mathbf{N}_{ij} = p_{ij} \mathbf{M}_i$. The condition $\mathbf{M}_i = \sum_j p'_{ji} \mathbf{M}'_j$ yields $p_i = \text{tr}[\rho \mathbf{M}_i] = \sum_j p'_{ji} \text{tr}[\rho \mathbf{M}'_j]$ thus showing that, instead of measuring \mathbf{M} , one can measure \mathbf{M}' in the same state ρ and then classically post-process the data by using the matrix (p'_{ij}) . Post-processing clean POVMs are free from this type of classical noise and it is easy to show that \mathbf{M} *is post-processing clean if and only if \mathbf{M} is of rank 1*¹³ (Theorem 3.4).
- (2a) \mathbf{M} *determines the past of the system* if it is *informationally complete*, i.e. the measurement outcome statistics $(p_i)_{i=1}^N$ determines the input state ρ , i.e. the condition $\text{tr}[\rho \mathbf{M}_i] = \text{tr}[\rho' \mathbf{M}_i]$ for all i implies that $\rho' = \rho$. Clearly, \mathbf{M} *determines the past of the system if and only if $N \geq d^2$ and any $d \times d$ -matrix B can be written as a linear combination of matrices \mathbf{M}_i , $i = 1, \dots, N^5$* (Prop. 18.1). We will construct later an in-

formationally complete (extreme) rank-1 POVM \mathbf{M} with the minimum number $N = d^2$ of outcomes. Generally, an informationally complete POVM need not be rank-1 but *if \mathbf{M} is informationally complete then its rank-1 refinement \mathbf{M}^1 is also informationally complete*³³.

(2b) \mathbf{M} is *extreme* if it is an extremal point of the convex set of all (discrete) POVMs of \mathcal{H} , i.e. $\mathbf{M} = \frac{1}{2}\mathbf{M}' + \frac{1}{2}\mathbf{M}''$ implies $\mathbf{M}' = \mathbf{M}$. Thus, extreme observables describe statistics of the pure quantum measurements, free from any classical randomness due to fluctuations in the measuring procedure (in the same way as pure states describe preparation procedures without classical randomness). One can show that *\mathbf{M} is extreme if and only if the matrices $|d_{ik}\rangle\langle d_{i\ell}|$, $i = 1, \dots, N$, $k, \ell = 1, \dots, m_i$, are linearly independent*³⁰ (Theorem 2.4),⁹ (Theorem 2). Especially, if \mathbf{M} is rank-1, i.e. $\mathbf{M}_i = |d_i\rangle\langle d_i|$, $d_i \neq 0$, then it is extreme if and only if the matrices \mathbf{M}_i are linearly independent. In this case, it is informationally complete if and only if $N = d^2$. Trivially, *if \mathbf{M} is extreme then its rank-1 refinement \mathbf{M}^1 is also extreme* implying, since $d^2 = N^1 = \sum_{i=1}^N m_i \geq N$ and $N \geq d^2$ imply $m_i \equiv 1$ and $N = d^2$, that *any extreme informationally completely POVM is necessarily of rank 1*. Finally, we note that PVMs are automatically extreme.

(3a) \mathbf{M} *determines its values x_i* if each \mathbf{M}_i is of (operator) norm 1, i.e. \mathbf{M}_i has the eigenvalue 1 (with the unit eigenvector φ_i). In this case, for any outcome x_j one can pick a state $\rho = |\varphi_j\rangle\langle\varphi_j|$ such that $p_i = \text{tr}[\rho\mathbf{M}_i] = \delta_{ij}$ for all i , i.e. the observable \mathbf{M} has the value x_j in the state ρ with probabilistic certainty. Note that this holds only in finite dimensions. Generally, a norm-1 (effect) operator can have a fully continuous spectrum (i.e. no eigenvalues at all). However, even in such a case, for each j and any $\varepsilon \in (0, 1)$, there is a state $\rho = |\varphi_j\rangle\langle\varphi_j|$ such that $\text{tr}[\rho\mathbf{M}_j] = 1 - \varepsilon$. Clearly, a rank-1 norm-1 POVM is a PVM and any PVM is of norm-1. This holds for discrete observables. As a counterexample, consider the canonical phase observable which is rank-1 norm-1 ‘continuous’ POVM but not projection valued.

(3b) \mathbf{M} is *pre-processing maximal (pre-processing clean)* if the condition $\mathbf{M}_i = \Phi(\mathbf{M}'_i)$ for all i (where $\Phi : \mathcal{M}_{d'}(\mathbb{C}) \rightarrow \mathcal{M}_d(\mathbb{C})$ is a Heisenberg channel and \mathbf{M}' is a POVM on the possibly different Hilbert space $\mathcal{H}' \cong \mathbb{C}^{d'}$ with $N' = N$ outcomes) implies that $\mathbf{M}'_i = \Theta(\mathbf{M}_i)$ for all i where $\Theta : \mathcal{M}_d(\mathbb{C}) \rightarrow \mathcal{M}_{d'}(\mathbb{C})$ is some Heisenberg channel. The condition $\mathbf{M}_i = \Phi(\mathbf{M}'_i)$ can be written in the form $p_i = \text{tr}[\rho\mathbf{M}_i] = \text{tr}[\rho\Phi(\mathbf{M}'_i)] = \text{tr}[\Phi_*(\rho)\mathbf{M}'_i]$ so

that to get the probabilities p_i one can equally well measure M' in the state $\Phi_*(\rho)$, i.e. M' is ‘better’ measurement in this sense and M is obtained from it by adding quantum noise in ρ (characterized by the channel Φ). Hence, pre-processing clean observables are free from this type of quantum noise. Since, using the Naimark dilation $(\mathcal{H}_\oplus, P, J)$ of M , $M_i = J^*P_iJ = \Phi(P_i)$, where Φ is the (rank-1) isometry channel $J^*(\cdot)J$, to show that M is pre-processing clean, one must find a channel Θ such that $P_i = \Theta(M_i)$ holds and thus, since $1 = \|P_i\| = \|\Theta(M_i)\| \leq \|\Theta\| \|M_i\| = \|M_i\| \leq 1$ implies $\|M_i\| = 1$, each M_i is of norm 1, i.e. M determines its values. Actually, Remark 1 shows that $M_i = \Phi_J(P'_i)$ must be connected to a *rank-1* PVM P' via some channel. We will show that, in finite dimensions, *pre-processing clean POVMs are exactly norm-1 POVMs* and exactly of the form $M_i = E_i \oplus F_i$ where E is a PVM ($E_i \neq 0$ for all i) and F a POVM ($\Theta(F_i) = 0$ for all i) acting on orthogonal subspaces of \mathcal{H} . Hence, for any pre-processing clean POVM M , there exists a projection (onto a subspace) such that the projected POVM E is projection valued. Especially, PVMs are pre-processing clean³¹.

We have seen that ‘optimal observables’ must be of rank 1, see (1a) and (1b) above. Moreover, observables satisfying (2a) or (2b) can be maximally refined into rank-1 observables which share the same optimality criteria as the original POVMs. If M is rank-1 then the map $\mathbb{C}^N \ni (c_1, \dots, c_2) \mapsto \sum_{i=1}^N c_i M_i$ is surjective iff (2a) holds and injective iff (2b) holds. We will construct a POVM for which all conditions (1a), (1b), (2a), and (2b) hold. PVMs are optimal observables in the sense of (3a) and (3b). In addition, the rank-1 refinement of a PVM is also projection valued (and of norm-1). However, it is easy to construct a norm-1 POVM whose rank-1 refinement is not of norm 1. For example, in \mathbb{C}^3 (with the basis $|0\rangle, |1\rangle, |2\rangle$), one can define 2-valued norm-1 POVM $M_1 = |1\rangle\langle 1| + \frac{1}{3}|0\rangle\langle 0|$, $M_2 = |2\rangle\langle 2| + \frac{2}{3}|0\rangle\langle 0|$ whose refinement M^1 has an effect $M_{12}^1 = \frac{1}{3}|0\rangle\langle 0|$ of norm $\frac{1}{3}$. Note that $M_i = E_i \oplus F_i$ where $E_1 = |1\rangle\langle 1|$, $E_2 = |2\rangle\langle 2|$ constitutes a PVM in a 2-dimensional space, and $F_1 = \frac{1}{3}|0\rangle\langle 0|$, $F_2 = \frac{2}{3}|0\rangle\langle 0|$.

To conclude, there are essentially two sorts of optimal observables: rank-1 PVMs and extreme informationally complete POVMs. Since they are extreme (2b) and rank-1, i.e. post processing clean (1b), they are free from classical noise due to the mixing of measurement schemes or data processing. Moreover, they determine the future of the system (1a). Since a pre-processing clean POVM has at most $N = d$ outcomes and an informationally complete POVM has at least $N = d^2$ outcomes, a pre-processing clean POVM (e.g. a PVM) cannot be

informationally complete and vice versa, i.e. an informationally complete POVM is never free from quantum noise. Moreover, *the determination of the past (2a) and the values (3a) are complementary properties*. However, when one assumes that only a restricted class of states (related to a subspace $\mathcal{H} \subseteq \mathcal{H}_\oplus$) can be determined completely then these complementary properties can be combined as follows:

One can pick a d^2 -outcome extreme informationally complete rank-1 POVM $M_i = |d_i\rangle\langle d_i|$, $i = 1, \dots, d^2$, and its minimal Naimark dilation with the rank-1 PVM $P_i = |e_i\rangle\langle e_i|$ acting in a d^2 -dimensional space $\mathcal{H}_\oplus = \mathcal{H} \oplus \mathcal{H}^\perp$ with the orthogonal basis $\{e_i\}_{i=1}^{d^2}$ (recall that $d = \dim \mathcal{H}$). Now, for any subsystem's state ρ , one gets $p_i = \langle d_i|\rho|d_i\rangle = \langle e_i|J\rho J^*|e_i\rangle$ and P_i is informationally complete only within the set of states of the subspace \mathcal{H} . Instead of measuring M one can prepare a state of $\mathcal{H} \cong J\mathcal{H}$ and then perform a measurement of P to get probabilities p_i and posterior states σ_i (see item (1a) above) since the nuclear instrument $\mathcal{I}_i(\rho) = \langle d_i|\rho|d_i\rangle\sigma_i$ implementing M can be trivially extended to an instrument $\bar{\mathcal{I}}$ of P via $\bar{\mathcal{I}}_i(\bar{\rho}) = \langle e_i|\bar{\rho}|e_i\rangle\sigma_i$ where $\bar{\rho}$ is a state of \mathcal{H}_\oplus . Below we study sequential and joint measurements of optimal POVMs with other observables.

Let $M = (M_i)_{i=1}^N$ and $M' = (M'_j)_{j=1}^{N'}$ be POVMs as in the beginning of this introduction, and let $\mathcal{I} = (\mathcal{I}_i)$ be an instrument implementing M (i.e. any $\mathcal{I}_i : \mathcal{M}_d(\mathbb{C}) \rightarrow \mathcal{M}_{d'}(\mathbb{C})$ is of the form $\mathcal{I}_i(\rho) = \sum_s A_{is}\rho A_{is}^*$ and $M_i = \sum_s A_{is}^* A_{is}$). Suppose then that one measures first M in the state ρ (described by \mathcal{I}) and then M' in the transformed (conditional) state $p_i^{-1}\mathcal{I}_i(\rho)$ if the outcome x_i is obtained (with the probability $p_i > 0$) in the first measurement of M . This sequential measurement can be described by a joint POVM $J = (J_{ij})$ where $J_{ij} = \mathcal{I}_i^*(M'_j)$ since the conditional probability is $\text{tr}[\rho J_{ij}] = \text{tr}[p_i^{-1}\mathcal{I}_i(\rho)M'_j] p_i$. Hence, *a sequential measurement of M and M' can be interpreted as a joint measurement of M and the disturbed POVM M'' , $M''_j = \sum_i J_{ij} = \Phi(M'_j)$ of the same Hilbert space (here $\Phi = \sum_i \mathcal{I}_i^*$ is the total Heisenberg channel of \mathcal{I}).*

Indeed, any POVMs $M = (M_i)$ and $M'' = (M''_j)$ (of the same Hilbert space $\mathcal{H}'' = \mathcal{H}$) are jointly measurable if there exists a POVM $N = (N_{ij})$ such that M and M'' are the margins of N , i.e., $M_i = \sum_{j=1}^{N''} N_{ij}$ and $M''_j = \sum_{i=1}^N N_{ij}$. If $(\mathcal{H}_\oplus, J, P)$ is a (minimal) Naimark dilation of M , since $N_{ij} \leq \sum_{j=1}^{N''} N_{ij} = J^*P_iJ$ implies the existence of $P_{ij} \leq P_i$ ($m_i \times m_i$ -identity matrix), it follows that $N_{ij} = J^*P_{ij}J$ where P_{ij} is a (unique) positive semidefinite $m_i \times m_i$ -matrix such that $\sum_j P_{ij} = P_i$. Hence, for each $i \leq N$, the map $j \mapsto P_{ij}$ is a POVM acting in the subspace $P_i\mathcal{H}_\oplus$ with N'' values so that there is a channel Φ_i such that $\Phi_i(P'_j) = P_{ij}$,

see Remark 1. Here $\mathbf{P}' = (\mathbf{P}'_j)_{j=1}^{N''}$ is a fixed rank-1 PVM acting in a minimal N'' -dimensional Hilbert space $\mathcal{H}_{N''} \cong \mathbb{C}^{N''}$. Indeed if, say, $\mathbf{M}''_k = 0$ then $\mathbf{N}_{ik} \leq \sum_{i'=1}^N \mathbf{N}_{i'k} = \mathbf{M}''_k$ yields $\mathbf{N}_{ik} = 0$ and thus $\mathbf{P}_{ik} = 0$ for all $i \leq N$. Hence, $\mathbf{P}'_k = 0$ and N'' is the number of nonzero effects \mathbf{M}''_j since usually we assume that $\mathbf{M}''_j \neq 0$ for all $j = 1, \dots, N''$. Define an instrument \mathcal{I} by $\mathcal{I}_i^*(B) = J^* \Phi_i(B) J$, $B \in \mathcal{M}_{N''}(\mathbb{C})$. Clearly, $\mathbf{N}_{ij} = \mathcal{I}_i^*(\mathbf{P}'_j)$ and, thus, any *joint measurement of \mathbf{M} and \mathbf{M}'' can be interpreted as a sequential measurement of \mathbf{M} followed by a rank-1 PVM \mathbf{P}'* . Note that $\mathbf{M}''_j = \Phi'(\mathbf{P}'_j)$ where $\Phi'(B) = J^* \sum_i \Phi_i(B) J$.

In conclusion, a sequential measurement of \mathbf{M} and \mathbf{M}' defines a joint observable \mathbf{J} with the margins \mathbf{M} and $\mathbf{M}'' = \Phi(\mathbf{M}')$. If we put $\mathbf{N} = \mathbf{J}$ above, we see that this measurement of \mathbf{J} can be interpreted as a new sequential measurement of \mathbf{M} and a rank-1 PVM \mathbf{P}' . In addition, $\mathbf{M}'' = \Phi'(\mathbf{P}')$. Thus, the latter observable in a sequential set-up can be assumed to be very optimal: free from both classical and quantum noise. Next we study how the optimality criteria (1a)—(3b) affect the joint measurability of an optimal observable with other observables.

1. If \mathbf{M} is rank-1 ($m_i \equiv 1$) then any \mathbf{P}_{ij} is a 1×1 -matrix, i.e. a number p_{ij} , and $\mathbf{N}_{ij} = p_{ij} \mathbf{M}_i$ is rank-1 (and a post-processing of \mathbf{M}). Since (p_{ij}) is a probability matrix, also \mathbf{M}'' is a smearing (post-processing) of \mathbf{M} , i.e. $\mathbf{M}''_j = \sum_i p_{ij} \mathbf{M}_i$, see item (1b) above. Moreover, the \mathbf{M} -compatible instrument \mathcal{I} is nuclear (1a) and $\mathbf{N}_{ij} = \mathbf{J}_{ij} = \text{tr}[\sigma_i \mathbf{M}'_j] \mathbf{M}_i$ from where one can read the transition probabilities $p_{ij} = \text{tr}[\sigma_i \mathbf{M}'_j]$ (where the states σ_i determine \mathcal{I} completely) showing that, if one gets x_i in the first \mathbf{M} -measurement, then the instrument ‘prepares’ the post-measurement state σ_i which is the input state for the second \mathbf{M}' -measurement giving the probability distribution p_{ij} , $j = 1, \dots, N'$, and $\text{tr}[\rho \mathbf{J}_{ij}] = p_i p_{ij}$ where $p_i = \text{tr}[\rho \mathbf{M}_i]$. Hence, after a measurement of a rank-1 POVM there is no need to perform any extra measurements to get more information. It should be stressed that, even if $\mathbf{M}''_j = \Phi'(\mathbf{P}'_j)$, the channel Φ' adds so much quantum noise to the rank-1 PVM \mathbf{P}' that it becomes the fuzzy version \mathbf{M}'' of \mathbf{M} ³². We easily see that Φ is entanglement braking since it is associated with a nuclear instrument \mathcal{I}' of \mathbf{M} : Since $\mathbf{P}'_j = |e_j\rangle\langle e_j|$ one can define states $\sigma'_i = \sum_j p_{ij} |e_j\rangle\langle e_j|$ and a nuclear instrument $\mathcal{I}'_i(\rho) = \text{tr}[\rho \mathbf{M}_i] \sigma'_i$ (or $\mathcal{I}'_i^*(B) = \text{tr}[B \sigma'_i] \mathbf{M}_i$) such that $\mathcal{I}'_i^*(\mathbf{P}'_j) = p_{ij} \mathbf{M}_i = \mathbf{N}_{ij}$ and $\Phi'(B) = \sum_i \mathcal{I}'_i^*(B) = \sum_i \text{tr}[B \sigma'_i] \mathbf{M}_i$. Hence, in this sequential measurement, the latter observable \mathbf{M}'' , which arises as the second marginal of the joint observable \mathbf{J} ,

is obtained both through adding classical noise to the observable M first measured (i.e., as a post-processing $M''_j = \sum_i p_{ij} M_i$) and through adding quantum noise to the observable M' actually measured after M in the form of the pre-processing $M''_j = \Phi(M'_j) = \sum_i \text{tr}[\sigma_i M'_j] M_i$. The same results naturally apply in the situation where we modify the measurement of the first observable M and measure some rank-1 PVM P' after it to obtain M'' as the second marginal. In this case M'' is a classical smearing of the rank-1 M and a quantum smearing of the rank-1 PVM P' .

2. If M is informationally complete (2a) then $N = J$ is also informationally complete: having access to probabilities $\text{tr}[\rho N_{ij}]$, one can solve ρ from the probabilities $p_i = \text{tr}[\rho M_i] = \sum_{j=1}^{N''} \text{tr}[\rho N_{ij}]$. Hence, trivially, if already M determines the past, then its subsequent measurements cannot increase the (already maximal) state distinguishing power. Suppose now that $N_{ij} = J_{ij} = \mathcal{I}_i^*(M'_j)$ for some instrument \mathcal{I} measuring the informationally complete M and some subsequently measured M' giving rise to the second marginal $M''_j = \Phi(M'_j)$ for the total channel $\Phi = \sum_i \mathcal{I}_i^*$. If we also assume that M'' is informationally complete, i.e., we jointly measure two informationally complete observables in a sequential setting, then the Heisenberg channel Φ is surjective (in this finite-dimensional case) and the corresponding Schrödinger channel $\Phi_* = \sum_i \mathcal{I}_i$ is injective. Indeed, since M'' is informationally complete, the map $\mathcal{M}_d(\mathbb{C}) \ni \rho \mapsto (\text{tr}[\rho M''_j])_j$ is injective, and, since this map is the composition of the maps Φ_* and $\sigma \mapsto (\text{tr}[\sigma M'_j])_j$, as the first of these maps, Φ_* has to be injective.

If M is extreme (2b) then N is the unique joint POVM which has the margins M and M'' . This is proven as follows: Since, for $N_{ij} = J^* P_{ij} J$ and $N'_{ij} = J^* P'_{ij} J$, the condition $M''_j = \sum_{i=1}^N N_{ij} = \sum_{i=1}^N N'_{ij}$ can be written in the form $J^* D_j J = 0$ where $D_j = \bigoplus_{i=1}^N (P_{ij} - P'_{ij})$. If M is extreme then $D_j \equiv 0$, i.e. $P_{ij} \equiv P'_{ij}$, and $N = N'$. If, in addition, M is rank-1 then $N_{ij} = p_{ij} M_i$ and $M''_j = \sum_i p_{ij} M_i$ and we have the chain of bijections: $N \mapsto M'' \mapsto (p_{ij}) \mapsto N$.

3. If $M = P$ is a PVM, i.e. $P_k P_\ell \equiv \delta_{k\ell} P_k$, (and thus $\mathcal{H}_\oplus = \mathcal{H}$ and J is the identity map) then $N_{ij} = P_{ij}$ where each map $j \mapsto P_{ij}$ is a (subnormalized) POVM which commutes with P , i.e. $P_{ij} P_k \equiv P_k P_{ij}$, since $P_{ij} \leq P_i$. Thus, any M'' compatible with a PVM P commutes with P . If, moreover, P is of rank-1 then $P_{ij} = p_{ij} P_i$. Note that N (or M'') needs not to be a PVM or even of norm 1 (e.g. consider $P_1 = |1\rangle\langle 1|$,

$$P_2 = |2\rangle\langle 2| = P_{22}, P_{11} = P_{12} = \frac{1}{2}|1\rangle\langle 1| = M''_2, M''_1 = \frac{1}{2}|1\rangle\langle 1| + |2\rangle\langle 2| \text{ in } \mathbb{C}^2).$$

In this paper, we generalize the above results to the case of arbitrary observables (with sufficiently ‘nice’ value spaces) acting in separable Hilbert spaces. For example, consider the single-mode optical field with the Hilbert space \mathcal{H} spanned by the photon number states $|n\rangle$, $n = 0, 1, 2, \dots$, associated with the number operator $N = a^*a = \sum_{n=0}^{\infty} n|n\rangle\langle n|$ where $a = \sum_{n=0}^{\infty} \sqrt{n+1}|n\rangle\langle n+1|$. Define the position and momentum operators $Q = \frac{1}{\sqrt{2}}(a^* + a)$ and $P = \frac{i}{\sqrt{2}}(a^* - a)$ which, in the position representation, are the usual multiplication and differentiation operators, $(Q\psi)(x) = x\psi(x)$ and $(P\psi)(x) = -i d\psi(x)/dx$ (we set $\hbar = 1$). Define the Weyl operator (or the displacement operator of the complex plane) $D(z) = e^{za^* - \bar{z}a}$, $z \in \mathbb{C}$. Recall that Weyl operators are associated to a unitary representation of the Heisenberg group \mathbb{H} , or to a projective representation of the additive group $\mathbb{C} \cong \mathbb{R}^2$. Let $q, p \in \mathbb{R}$ and $z = (q + ip)/\sqrt{2}$. Then

$$D(q, p) = D(z) = e^{ipQ - iqP} = e^{-iqp/2} e^{ipQ} e^{-iqP} = e^{iqp/2} e^{-iqP} e^{ipQ},$$

i.e., for all $\psi \in \mathcal{H} \cong L^2(\mathbb{R})$, $(e^{ipQ}\psi)(x) = e^{ipx}\psi(x)$, $(e^{iqP}\psi)(x) = \psi(x + q)$, and

$$(D(q, p)\psi)(x) = e^{-iqp/2} e^{ipx}\psi(x - q).$$

One can measure the following physically relevant POVMs:

- Rotated quadrature operators $Q_\theta = (\cos \theta)Q + (\sin \theta)P$ where $\theta \in [0, 2\pi)$ so that $Q_0 = Q$ and $Q_{\pi/2} = P$. In the position representation, the spectral measure of Q is the canonical spectral measure, $[Q(X)\psi](x) = \chi_X(x)\psi(x)$, $X \subseteq \mathbb{R}$ (Borel set), so that the spectral measure (rank-1 PVM) of Q_θ is $Q_\theta(X) = R(\theta)Q(X)R(\theta)^*$ where $R(\theta) = e^{i\theta N} = \sum_{n=0}^{\infty} e^{in\theta}|n\rangle\langle n|$ is the (unitary) rotation operator. Rotated quadratures can be measured by a balanced homodyne detector where the phase shift θ is caused by a phase shifter. A single Q_θ cannot be informationally complete (as a PVM) but the whole measurement assemblage $\{Q_\theta(X)\}_{\substack{\theta \in [0, \pi) \\ X \subseteq \mathbb{R}}}$ forms an informationally complete set of effects. Actually, a rank-1 POVM $\mathbf{G}_{\text{ht}}(\Theta \times X) = \frac{1}{\pi} \int_{\Theta} Q_\theta(X) d\theta$ determines the input state completely (optical homodyne tomography, OHT). Note that $\|\mathbf{G}_{\text{ht}}(\Theta \times X)\| \leq \frac{1}{\pi} \int_{\Theta} d\theta < 1$ if $\Theta \subseteq [0, \pi)$ is not of ‘length’ π .
- The number operator $N = \sum_{n=0}^{\infty} n|n\rangle\langle n|$ whose spectral measure (rank-1 PVM) is $n \mapsto \mathbf{N}_n = |n\rangle\langle n|$ (an ideal photon detector with the 100 % efficiency).

- An unsharp (rank- ∞) number observable (POVM)

$$n \mapsto \mathbf{N}_n^\epsilon = \sum_{m=n}^{\infty} \binom{m}{n} \epsilon^n (1-\epsilon)^{m-n} |m\rangle\langle m|$$

(a nonideal photon detector with quantum efficiency $\epsilon \in [0, 1)$). Now \mathbf{N}^ϵ is neither of norm-1 nor informationally complete, since it is commutative⁴, and $\lim_{\epsilon \rightarrow 1} \mathbf{N}_n^\epsilon = \mathbf{N}_n$ (by $0^0 = 1$).

- Covariant phase space observables (POVMs)

$$\mathbf{G}_S(Z) = \frac{1}{\pi} \int_Z D(z) S D(z)^* d^2z = \frac{1}{2\pi} \int_Z D(q, p) S D(q, p)^* dq dp, \quad Z \subseteq \mathbb{C},$$

where S is (essentially) the reference state of an eight-port (or double) homodyne detector. In practice \mathbf{G}_S can be viewed as a joint measurement of unsharp rotated quadratures (e.g., unsharp position and momentum). Moreover, \mathbf{G}_S is rank-1 if and only if $S = |\psi\rangle\langle\psi|$ where ψ is a unit vector. Note that $\|\mathbf{G}_S(Z)\| \leq \frac{1}{\pi} \int_Z d^2z < 1$ when the area of the set Z is small enough.

- Covariant phase observables (POVMs)

$$\Phi_C(\Theta) = \frac{1}{2\pi} \int_{\Theta} R(\theta) C R(\theta)^* d\theta = \sum_{n,m=0}^{\infty} C_{nm} \int_{\Theta} e^{i(n-m)\theta} \frac{d\theta}{2\pi} |n\rangle\langle m|, \quad \Theta \subseteq [0, 2\pi),$$

where $C = \sum_{n,m=0}^{\infty} C_{nm} |n\rangle\langle m|$ is a positive sesquilinear form with the unit diagonal ($C_{nn} \equiv 1$), i.e. a phase matrix. If $C_{nm} \equiv 1$ we get the canonical phase observable Φ_{can} ,

$$\Phi_{\text{can}}(\Theta) = \sum_{n,m=0}^{\infty} \int_{\Theta} e^{i(n-m)\theta} \frac{d\theta}{2\pi} |n\rangle\langle m|, \quad \Theta \subseteq [0, 2\pi),$$

whereas the angle margin of \mathbf{G}_S is called a phase space phase observable. Both can be measured by double homodyne detection³⁶. Any phase observable is never projection valued and is a preprocessed version of the canonical phase given by a (Schur type) quantum channel. Note that Φ_{can} is not informationally complete (consider number states) but it is norm-1. However, Φ_{can} is not pre-processing clean since its nontrivial effects cannot have eigenvalues⁵ (Theorem 8.2).

A. Definitions and mathematical background

In this paper, $\mathbb{N} = \{1, 2, \dots\}$, i.e., 0 is not included in the set of natural numbers. We define an empty sum $\sum_{j=1}^0 (\dots)$ to be equal to zero. When \mathcal{H} is a Hilbert space, we denote

by $\mathcal{L}(\mathcal{H})$ the algebra of bounded linear operators on \mathcal{H} and by $I_{\mathcal{H}}$ the unit element of this algebra (the identity operator on \mathcal{H}); by ‘Hilbert space’ we always mean a complex Hilbert space. The inner product of any Hilbert space will be simply denoted by $\langle \cdot | \cdot \rangle$ since the Hilbert space in question should always be clear from the context, and the inner product is chosen to be linear in the second argument. By $\mathcal{P}(\mathcal{H})$, we denote the set of projections of \mathcal{H} , i.e., operators $P \in \mathcal{L}(\mathcal{H})$ such that $P = P^* = P^2$. An operator $E \in \mathcal{L}(\mathcal{H})$ is called effect if $0 \leq E \leq I_{\mathcal{H}}$ holds. Especially, any projection is an effect, the so-called sharp effect. We let $\mathcal{T}(\mathcal{H})$ stand for the set of trace-class operators on \mathcal{H} , i.e., $\text{tr}[|T|] < \infty$ for all $T \in \mathcal{T}(\mathcal{H})$. We denote the set of positive trace-1 operators in $\mathcal{T}(\mathcal{H})$ by $\mathcal{S}(\mathcal{H})$; in quantum physics, these normalized positive states of $\mathcal{L}(\mathcal{H})$ are identified with the physical states of the system described by \mathcal{H} . Note that $\mathcal{P}(\mathcal{H}) \cap \mathcal{S}(\mathcal{H})$ consists of rank-1 projections $|\psi\rangle\langle\psi|$ (where $\psi \in \mathcal{H}$ is a unit vector).

When μ and ν are positive measures on a measurable space (Ω, Σ) (where $\Omega \neq \emptyset$ is a set and Σ is a σ -algebra of subsets of Ω) we say that μ is absolutely continuous with respect to ν and denote $\mu \ll \nu$ if $\mu(X) = 0$ whenever $\nu(X) = 0$. When both $\mu \ll \nu$ and $\nu \ll \mu$, we denote $\mu \sim \nu$ and say that μ and ν are equivalent.

Let (Ω, Σ) be a measurable space and \mathcal{H} be a Hilbert space. A map $\mathbf{M} : \Sigma \rightarrow \mathcal{L}(\mathcal{H})$ is said to be a *normalized positive-operator-valued measure (POVM)* if, for all $\rho \in \mathcal{S}(\mathcal{H})$, the set function $X \mapsto \text{tr}[\rho\mathbf{M}(X)]$, denoted hereafter by $p_{\rho}^{\mathbf{M}}$, is a probability measure or, equivalently, $\mathbf{M}(X) \geq 0$ for all $X \in \Sigma$, $\mathbf{M}(\Omega) = I_{\mathcal{H}}$, and, for any pairwise disjoint sequence $X_1, X_2, \dots \in \Sigma$, one has $\mathbf{M}(\cup_j X_j) = \sum_j \mathbf{M}(X_j)$ (ultra)weakly. Denote the set of POVMs from Σ to $\mathcal{L}(\mathcal{H})$ by $\text{Obs}(\Sigma, \mathcal{H})$. When $\mathbf{P}(X) \in \mathcal{P}(\mathcal{H})$ for all $X \in \Sigma$ for a POVM $\mathbf{P} : \Sigma \rightarrow \mathcal{L}(\mathcal{H})$, we say that \mathbf{P} is a *normalized projection-valued measure (PVM)* or a spectral measure. We extend the notions of absolute continuity and equivalence introduced above for scalar measures in the obvious way and thus may write, e.g., for a POVM $\mathbf{M} \in \text{Obs}(\Sigma, \mathcal{H})$ and a measure $\mu : \Sigma \rightarrow [0, \infty]$, $\mathbf{M} \ll \mu$ if $\mathbf{M}(X) = 0$ whenever $\mu(X) = 0$ and, for another POVM $\mathbf{N} : \Sigma \rightarrow \mathcal{L}(\mathcal{K})$, where \mathcal{K} is some Hilbert space, $\mathbf{M} \ll \mathbf{N}$ if $\mathbf{N}(X) = 0$ implies $\mathbf{M}(X) = 0$. We say that \mathbf{M} is *discrete* if there exist distinct points $\{x_i\}_{i=1}^N \subseteq \Omega$, $N \in \mathbb{N} \cup \{\infty\}$, and effects $\{\mathbf{M}_i\}_{i=1}^N \subseteq \mathcal{L}(\mathcal{H})$ such that $\mathbf{M} = \sum_{i=1}^N \mathbf{M}_i \delta_{x_i}$ where δ_x is a Dirac (point) measure concentrated on x . Now $\mathbf{M} \ll \sum_{i=1}^N \delta_{x_i}$. A discrete observable \mathbf{M} can naturally be identified with the effects \mathbf{M}_i and we will use the notation $(\mathbf{M}_i)_{i=1}^N$ for \mathbf{M} if the outcomes $x \in \Omega$ are not relevant. Note that, if \mathcal{H} is separable, and $\mathbf{M} \in \text{Obs}(\Sigma, \mathcal{H})$, picking any state $\rho \in \mathcal{S}(\mathcal{H})$ which is faithful, i.e., $\text{tr}[\rho A] = 0$ implies

$A = 0$ for any positive $A \in \mathcal{L}(\mathcal{H})$ (or, equivalently, the kernel of ρ is $\{0\}$), we have $\mathbf{M} \sim p_\rho^{\mathbf{M}}$.

In quantum physics, POVMs are associated in a one-to-one fashion with observables of the system. The observables associated with PVMs are called sharp. In this view, the number $p_\rho^{\mathbf{M}}(X) = \text{tr}[\rho \mathbf{M}(X)] \in [0, 1]$ is the probability of obtaining a value within the outcome set $X \in \Sigma$ when measuring $\mathbf{M} \in \text{Obs}(\Sigma, \mathcal{H})$ and the system being measured is in the quantum state $\rho \in \mathcal{S}(\mathcal{H})$. In realistic physical experiments, we measure only discrete observables which in many cases can be thought as discretizations of continuous observables, i.e. for any $\mathbf{M} \in \text{Obs}(\Sigma, \mathcal{H})$ one can choose pairwise disjoint sets $X_i \in \Sigma$ whose union is the whole Ω and define a discrete POVM by $\mathbf{M}_i = \mathbf{M}(X_i)$ (with the outcome set $\{1, \dots, N\}$ or \mathbb{N}). In this case, one can replace Σ with the sub- σ -algebra generated by the sets X_i .

Let \mathcal{A} and \mathcal{B} be C^* -algebras. We say that a linear map $\Phi : \mathcal{A} \rightarrow \mathcal{B}$ is n -positive ($n \in \mathbb{N}$) if the map

$$\mathcal{M}_n(\mathcal{A}) \ni (a_{ij})_{i,j=1}^n \mapsto (\Phi(a_{ij}))_{i,j=1}^n \in \mathcal{M}_n(\mathcal{B})$$

defined between the $n \times n$ -matrix algebras over the input and output algebras is positive. If Φ is n -positive for all $n \in \mathbb{N}$, Φ is said to be *completely positive*. Suppose that \mathcal{A} and \mathcal{B} are unital (with units $1_{\mathcal{A}}$ and $1_{\mathcal{B}}$) in which case Φ is called *unital* if $\Phi(1_{\mathcal{A}}) = 1_{\mathcal{B}}$. For any unital 2-positive map $\Phi : \mathcal{A} \rightarrow \mathcal{B}$ one has the Schwarz inequality, $\Phi(a)^* \Phi(a) \leq \Phi(a^* a)$ for all $a \in \mathcal{A}$. Suppose that \mathcal{A} and \mathcal{B} are von Neumann algebras. We say that a positive map $\Phi : \mathcal{A} \rightarrow \mathcal{B}$ is *normal*, if for any increasing (equivalently, decreasing) net $(a_\lambda)_\lambda \subseteq \mathcal{A}$ of self-adjoint operators, one has

$$\sup_\lambda \Phi(a_\lambda) = \Phi\left(\sup_\lambda a_\lambda\right),$$

where $\sup_\lambda a_\lambda$ is the supremum (ultraweak limit) of the increasing net (equivalently, with \sup replaced by \inf , the infimum, in the case of a decreasing net).

Fix a C^* -algebra \mathcal{A} and a Hilbert space \mathcal{H} . For any completely positive map $\Phi : \mathcal{A} \rightarrow \mathcal{L}(\mathcal{H})$, there is a Hilbert space \mathcal{M} , a unital $*$ -representation $\pi : \mathcal{A} \rightarrow \mathcal{L}(\mathcal{M})$, and an isometry $J : \mathcal{H} \rightarrow \mathcal{M}$ such that $\Phi(a) = J^* \pi(a) J$ for all $a \in \mathcal{A}$ and the linear hull of vectors $\pi(a) J \varphi$, $a \in \mathcal{A}$, $\varphi \in \mathcal{H}$, forms a dense subspace of \mathcal{M} . Such a triple (\mathcal{M}, π, J) is called as a *minimal Stinespring dilation* for Φ and it is unique up to unitary equivalence, i.e., if (\mathcal{M}', π', J') is another minimal dilation, then there is a unitary operator $U : \mathcal{M} \rightarrow \mathcal{M}'$ such that $U \pi(a) = \pi'(a) U$ for all $a \in \mathcal{A}$ and $U J = J'$.

Let \mathcal{H} and \mathcal{K} be Hilbert spaces. We call normal completely positive maps $\Phi : \mathcal{L}(\mathcal{K}) \rightarrow \mathcal{L}(\mathcal{H})$ satisfying $\Phi(I_{\mathcal{K}}) \leq I_{\mathcal{H}}$ as *operations*. When Φ is in addition unital, i.e., $\Phi(I_{\mathcal{K}}) = I_{\mathcal{H}}$, we call Φ as a *channel*. For a channel $\Phi : \mathcal{L}(\mathcal{K}) \rightarrow \mathcal{L}(\mathcal{H})$ one always has a minimal Stinespring dilation (\mathcal{K}', V) where \mathcal{K}' is a Hilbert space (which is separable if \mathcal{H} and \mathcal{K} are separable) and $V : \mathcal{H} \rightarrow \mathcal{K} \otimes \mathcal{K}'$ is an isometry such that $\Phi(B) = V^*(B \otimes I_{\mathcal{K}'})V$ for any $B \in \mathcal{L}(\mathcal{K})$ and vectors $(B \otimes I_{\mathcal{K}'})V\varphi$, $B \in \mathcal{L}(\mathcal{K})$, $\varphi \in \mathcal{H}$, span a dense subspace of $\mathcal{K} \otimes \mathcal{K}'$. For any normal linear map $\Phi : \mathcal{L}(\mathcal{K}) \rightarrow \mathcal{L}(\mathcal{H})$ there exists a (unique) predual map $\Phi_* : \mathcal{T}(\mathcal{H}) \rightarrow \mathcal{T}(\mathcal{K})$ such that

$$\mathrm{tr}[\Phi_*(T)A] = \mathrm{tr}[T\Phi(A)], \quad T \in \mathcal{T}(\mathcal{H}), \quad A \in \mathcal{L}(\mathcal{K}).$$

The version $\Phi : \mathcal{L}(\mathcal{K}) \rightarrow \mathcal{L}(\mathcal{H})$ is said to be in the *Heisenberg picture* and the version $\Phi_* : \mathcal{T}(\mathcal{H}) \rightarrow \mathcal{T}(\mathcal{K})$ is said to be in the *Schrödinger picture*. For a channel Φ , the Schrödinger channel Φ_* , when restricted onto $\mathcal{S}(\mathcal{H})$, describes how the system associated with \mathcal{H} transforms under Φ into another system associated with \mathcal{K} .

Let (Ω, Σ) be a measurable space and \mathcal{H} and \mathcal{K} Hilbert spaces. We say that a map $\mathcal{J} : \mathcal{L}(\mathcal{K}) \times \Sigma \rightarrow \mathcal{L}(\mathcal{H})$ is a (Heisenberg) *instrument* if

- (i) $\mathcal{J}(\cdot, X) : \mathcal{L}(\mathcal{K}) \rightarrow \mathcal{L}(\mathcal{H})$ is an operation for all $X \in \Sigma$,
- (ii) $\mathcal{J}(\cdot, \Omega)$ is a channel, and
- (iii) for any pairwise disjoint sequence $X_1, X_2, \dots \in \Sigma$ and any $A \in \mathcal{L}(\mathcal{K})$, $\mathcal{J}(A, \cup_j X_j) = \sum_j \mathcal{J}(A, X_j)$ (ultra)weakly.

For any instrument $\mathcal{J} : \mathcal{L}(\mathcal{K}) \times \Sigma \rightarrow \mathcal{L}(\mathcal{H})$, we define the predual (Schrödinger) instrument $\mathcal{J}_* : \mathcal{T}(\mathcal{H}) \times \Sigma \rightarrow \mathcal{T}(\mathcal{K})$,

$$\mathcal{J}_*(T, X) = [\mathcal{J}(\cdot, X)_*](T), \quad T \in \mathcal{T}(\mathcal{H}), \quad X \in \Sigma.$$

Note that, for an instrument \mathcal{J} , the map $\mathcal{J}(I_{\mathcal{K}}, \cdot) : \Sigma \rightarrow \mathcal{L}(\mathcal{H})$ is a POVM. On the other hand, for any $M \in \mathrm{Obs}(\Sigma, \mathcal{H})$ and a Hilbert space \mathcal{K} there is an instrument $\mathcal{J} : \mathcal{L}(\mathcal{K}) \times \Sigma \rightarrow \mathcal{L}(\mathcal{H})$ such that $M(X) = \mathcal{J}(I_{\mathcal{K}}, X)$ for all $X \in \Sigma$, i.e., $p_{\rho}^M = \mathrm{tr}[\mathcal{J}_*(\rho, \cdot)]$; we call \mathcal{J} an *M-instrument*. In a measurement of an observable associated with a POVM M , the system transforms conditioned by registering an outcome $x \in X$. This conditional state transformation is given by the operation $\mathcal{J}_*(\cdot, X)$ where \mathcal{J} is an *M-instrument*. The operator $\mathcal{J}_*(\rho, X)$ is a subnormalized state whose trace coincides with the probability $p_{\rho}^M(X)$

of registering an outcome in X . If $p_\rho^M(X) > 0$ then $[p_\rho^M(X)]^{-1} \mathcal{J}_*(\rho, X)$ is the corresponding conditional state.

II. GENERAL STRUCTURE OF A QUANTUM OBSERVABLE

In this section, we analyse the structure of an observable with a general value space on a system described by a separable Hilbert space. We will refer to the results reviewed in this section several times on the course of this paper.

Suppose that \mathcal{H} is a separable Hilbert space and let $\mathbf{h} = \{h_n\}_{n=1}^{\dim \mathcal{H}}$ be an orthonormal (ON) basis of \mathcal{H} and

$$V_{\mathbf{h}} := \text{lin}_{\mathbb{C}}\{h_n \mid 1 \leq n < \dim \mathcal{H} + 1\}.$$

Note that $V_{\mathbf{h}}$ is dense in \mathcal{H} . Let $V_{\mathbf{h}}^\times$ be the algebraic antidual of the vector space $V_{\mathbf{h}}$, that is, $V_{\mathbf{h}}^\times$ is the linear space consisting of all antilinear functions $c : V_{\mathbf{h}} \rightarrow \mathbb{C}$ (antilinearity means that $c(\alpha\psi + \beta\varphi) = \bar{\alpha}c(\psi) + \bar{\beta}c(\varphi)$ for all $\alpha, \beta \in \mathbb{C}$ and $\psi, \varphi \in V_{\mathbf{h}}$). By denoting $c_n = c(h_n)$ one sees that $V_{\mathbf{h}}^\times$ can be identified with the linear space of formal series $c = \sum_{n=1}^{\dim \mathcal{H}} c_n h_n$ where c_n 's are arbitrary complex numbers. Hence, $V_{\mathbf{h}} \subseteq \mathcal{H} \subseteq V_{\mathbf{h}}^\times$. Denote the dual pairing $\langle \psi | c \rangle := c(\psi) = \sum_{n=1}^{\dim \mathcal{H}} \langle \psi | h_n \rangle c_n$ and $\langle c | \psi \rangle := \overline{\langle \psi | c \rangle}$ for all $\psi \in V_{\mathbf{h}}$ and $c \in V_{\mathbf{h}}^\times$. Especially, $c_n = \langle h_n | c \rangle$. We say that a mapping $c : \Omega \rightarrow V_{\mathbf{h}}^\times$, $x \mapsto \sum_{n=1}^{\dim \mathcal{H}} c_n(x) h_n$ is (*weak*-*)*measurable* if its components $x \mapsto c_n(x)$ are measurable²³. Note that, if $c : \Omega \rightarrow \mathcal{H} \subseteq V_{\mathbf{h}}^\times$ is weak*-measurable then the maps $x \mapsto \langle \psi | c(x) \rangle$ are measurable for all $\psi \in \mathcal{H}$.

Let (Ω, Σ) be a measurable space and \mathcal{H}_\oplus denote a direct integral $\int_\Omega^\oplus \mathcal{H}(x) d\mu(x)$ of *separable* Hilbert spaces $\mathcal{H}(x)$ such that $\dim \mathcal{H}(x) = m(x) \in \mathbb{N} \cup \{0, \infty\}$; here μ is a σ -finite nonnegative measure on (Ω, Σ) ¹². Note that μ can be a probability measure everywhere in this paper; any σ -finite measure is equivalent with a probability measure. For each $f \in L^\infty(\mu)$, we denote briefly by \hat{f} the multiplicative (i.e. diagonalizable) bounded operator $(\hat{f}\psi)(x) := f(x)\psi(x)$ on \mathcal{H}_\oplus . Especially, one has the *canonical spectral measure* $\Sigma \ni X \mapsto \mathbf{P}_\oplus(X) := \hat{\chi}_X \in \mathcal{L}(\mathcal{H}_\oplus)$ (where χ_X is the characteristic function of $X \in \Sigma$). We say that an operator $D \in \mathcal{L}(\mathcal{H}_\oplus)$ is *decomposable* if there is a weakly μ -measurable field of operators $\Omega \ni x \mapsto D(x) \in \mathcal{L}(\mathcal{H}(x))$ such that $(D\psi)(x) = D(x)\psi(x)$ for all $\psi \in \mathcal{H}_\oplus$ and μ -a.a $x \in \Omega$; it is often denoted

$$D = \int_\Omega^\oplus D(x) d\mu(x).$$

We have the following theorem proved in^{23,31}:

Theorem 1. Let $\mathbf{M} : \Sigma \rightarrow \mathcal{L}(\mathcal{H})$ be a POVM and $\mu : \Sigma \rightarrow [0, \infty]$ a σ -finite measure such that $\mathbf{M} \ll \mu$. Let \mathbf{h} be an ON basis of \mathcal{H} . There exists a direct integral $\mathcal{H}_\oplus = \int_\Omega^\oplus \mathcal{H}(x) d\mu(x)$ (with $m(x) \leq \dim \mathcal{H}$) such that, for all $X \in \Sigma$,

(i) $\mathbf{M}(X) = J_\oplus^* \mathbf{P}_\oplus(X) J_\oplus$ where $J_\oplus : \mathcal{H} \rightarrow \mathcal{H}_\oplus$ is a linear isometry such that the set of linear combinations of vectors $\mathbf{P}_\oplus(X') J_\oplus \varphi$, $X' \in \Sigma$, $\varphi \in \mathcal{H}$, is dense in \mathcal{H}_\oplus .

(ii) There are measurable maps $d_k : \Omega \rightarrow V_{\mathbf{h}}^\times$ such that, for all $x \in \Omega$, the vectors $d_k(x) \neq 0$, $k < m(x) + 1$, are linearly independent, and

$$\langle \varphi | \mathbf{M}(X) \psi \rangle = \int_X \sum_{k=1}^{m(x)} \langle \varphi | d_k(x) \rangle \langle d_k(x) | \psi \rangle d\mu(x), \quad \varphi, \psi \in V_{\mathbf{h}},$$

(a minimal diagonalization of \mathbf{M}). In addition, there exist measurable maps $\Omega \ni x \mapsto g_\ell(x) \in V_{\mathbf{h}}$ such that $\langle d_k(x) | g_\ell(x) \rangle = \delta_{k\ell}$ (the Kronecker delta).

(iii) \mathbf{M} is a spectral measure if and only if J_\oplus is a unitary operator and thus \mathcal{H}_\oplus can be identified with \mathcal{H} .

A minimal Stinespring dilation for a POVM $\mathbf{M} : \Sigma \rightarrow \mathcal{L}(\mathcal{H})$ (viewed as a completely positive map $L^\infty(\mu) \rightarrow \mathcal{L}(\mathcal{H})$, $f \mapsto \int f d\mathbf{M}$, where μ is a probability measure such that $\mathbf{M} \ll \mu$) is called as a minimal Naimark dilation and it consists of a Hilbert space \mathcal{M} , an isometry $J : \mathcal{H} \rightarrow \mathcal{M}$, and a spectral measure $\mathbf{P} : \Sigma \rightarrow \mathcal{L}(\mathcal{M})$ such that $\mathbf{M}(X) = J^* \mathbf{P}(X) J$ and the vectors $\mathbf{P}(X) J \varphi$, $X \in \Sigma$, $\varphi \in \mathcal{H}$, span a dense subspace of \mathcal{M} . The above theorem tells that, whenever \mathcal{H} is separable, one can choose $\mathcal{M} = \mathcal{H}_\oplus = \int_\Omega^\oplus \mathcal{H}(x) d\mu(x)$ and \mathbf{P} to be the canonical spectral measure \mathbf{P}_\oplus .

A. Physical outcome spaces

It is reasonable to assume that a physically relevant outcome space (Ω, Σ) of an observable is regular or ‘nice’ enough. One can often suppose that Σ is *countably generated*, i.e. there exists a countable $S \subseteq \Sigma$ such that Σ is the smallest σ -algebra of Ω containing S . We will always consider any topological space T as a measurable space $(T, \mathcal{B}(T))$ where $\mathcal{B}(T)$ is the Borel σ -algebra of T . Furthermore, we equip any subset S of T with its subspace topology and the corresponding Borel σ -algebra $\mathcal{B}(S) = \mathcal{B}(T) \cap S$. We have the following proposition³⁷ (Proposition 3.2):

Proposition 1. *A measurable space (Ω, Σ) is countably generated if and only if there exists a map $f : \Omega \rightarrow \mathbb{R}$ such that*

(i) *for all $Y \in \mathcal{B}(\mathbb{R})$ the preimage $f^{-1}(Y) \in \Sigma$ (measurability) and*

(ii) *for all $X \in \Sigma$ there is $Y \in \mathcal{B}(\mathbb{R})$ such that $f^{-1}(Y) = X$.*

Recall that f satisfying (i) and (ii) is called exactly measurable. If (Ω, Σ) is countably generated and μ any σ -finite positive measure on Σ then $L^2(\mu)$ and $\mathcal{H}_\oplus = \int_\Omega^\oplus \mathcal{H}(x) d\mu(x)$ are separable.

We say that (Ω, Σ) is *nice* if it is countably generated and $f : \Omega \rightarrow \mathbb{R}$ of the above proposition meets the additional condition

(iii) $f(\Omega) \in \mathcal{B}(\mathbb{R})$.

In³⁷, nice spaces correspond to type \mathcal{B} -spaces (where \mathcal{B} refers to the notational conventions of the reference). Note that in this case actually $f(X) \in \mathcal{B}(\mathbb{R})$ for all $X \in \Sigma$ ³⁷ (Lemma 4.1). If, in addition, f is injective then the nice space (Ω, Σ) is a *standard Borel space* showing that nice spaces are generalizations of standard Borel spaces. Any Borel subset of a separable complete metric space is a standard Borel space and, indeed, any standard Borel space is σ -isomorphic to such a set or even to some compact metric space. Usually in physics, outcome spaces are finite-dimensional second countable Hausdorff manifolds which are (as locally compact spaces) standard Borel.

One can think of nice spaces as standard Borel spaces without the separability property (recall that Σ is separable if $\{x\} \in \Sigma$ for all $x \in \Omega$). For any $x \in \Omega$ one can define an atom $A_x := \bigcap \{X \in \Sigma \mid x \in X\} = f^{-1}(\{f(x)\})$ ³⁷ (Lemma 3.1) so that a nice space is standard Borel if and only if $A_x = \{x\}$ for all $x \in \Omega$. Hence, atoms of nice spaces may have an ‘inner structure’ (compare to the case of real world atoms).

Suppose that (Ω, Σ) is nice with an f satisfying (i)–(iii). Hence, $f(\Omega) \in \mathcal{B}(\mathbb{R})$ is a standard Borel space and, since two standard Borel spaces are σ -isomorphic if and only if they have the same cardinality, without restricting generality, we can assume that

- $f(\Omega) = \{1, 2, \dots, N\}$, $N \in \mathbb{N}$, or $f(\Omega) = \mathbb{N}$ (discrete case), or
- $f(\Omega) = \mathbb{R}$ (continuous case).

In the discrete case, we say that (Ω, Σ) is *discrete* and denote $X_i = f^{-1}(\{i\}) = A_{x_i}$, $i = 1, 2, \dots$, so that $X_i \cap X_j = \emptyset$, $i \neq j$, thus showing that Σ is the set of all unions of sets X_i and the empty set \emptyset . Moreover, any observable $\mathbf{M} : \Sigma \rightarrow \mathcal{L}(\mathcal{H})$ is discrete and, as earlier, can be identified with $(\mathbf{M}_i)_{i=1}^N$ where $\mathbf{M}_i = \mathbf{M}(X_i)$.

III. JOINT MEASURABILITY AND SEQUENTIAL MEASUREMENTS

If quantum devices can be applied simultaneously on the same system, we say that they are compatible. Simultaneously measurable observables are called jointly measurable. Let us give formal definitions for these notions.

Definition 1. Observables $\mathbf{M}_i : \Sigma_i \rightarrow \mathcal{L}(\mathcal{H})$ with outcome spaces (Ω_i, Σ_i) , $i = 1, 2$, are *jointly measurable* if they are *margins* of a *joint observable* $\mathbf{N} : \Sigma_1 \otimes \Sigma_2 \rightarrow \mathcal{H}$ defined on the product σ -algebra $\Sigma_1 \otimes \Sigma_2$ (generated by sets $X \times Y$, $X \in \Sigma_1$, $Y \in \Sigma_2$), i.e.,

$$\mathbf{M}_1(X) = \mathbf{N}(X \times \Omega_2), \quad \mathbf{M}_2(Y) = \mathbf{N}(\Omega_1 \times Y), \quad X \in \Sigma_1, \quad Y \in \Sigma_2.$$

Especially, \mathbf{M}_1 and \mathbf{M}_2 are jointly measurable if (and only if) they are *functions* or *relabelings* of a third observable $\mathbf{M} \in \text{Obs}(\Sigma, \mathcal{H})$, i.e. for both $i = 1, 2$ one has $\mathbf{M}_i(X_i) = \mathbf{M}(f_i^{-1}(X_i))$ for all $X_i \in \Sigma_i$ where $f_i : \Omega_i \rightarrow \Omega$ is a measurable function. Now a joint observable \mathbf{N} is defined by $\mathbf{N}(X \times Y) = \mathbf{M}(f_1^{-1}(X) \cap f_2^{-1}(Y))$ for all $X \in \Sigma_1$ and $Y \in \Sigma_2$. Note that, in this case, all the three measurable spaces can be arbitrary⁵ (Chapter 11). This implies that, in particular, any observable is jointly measurable with its relabelings.

Definition 2. Similarly, we say that an observable $\mathbf{M} : \Sigma \rightarrow \mathcal{L}(\mathcal{H})$ and a channel $\Phi : \mathcal{L}(\mathcal{K}) \rightarrow \mathcal{L}(\mathcal{H})$ are *compatible* if there exists a *joint instrument* $\mathcal{J} : \mathcal{L}(\mathcal{K}) \times \Sigma \rightarrow \mathcal{L}(\mathcal{H})$ such that

$$\mathbf{M}(X) = \mathcal{J}(I_{\mathcal{K}}, X), \quad \Phi(B) = \mathcal{J}(B, \Omega), \quad X \in \Sigma, \quad B \in \mathcal{L}(\mathcal{K}).$$

In this case, we also say that Φ is an *M-channel* and \mathcal{J} is an *M-instrument*.

The above means, when \mathbf{M} and Φ are compatible, there exists a measurement of \mathbf{M} such that Φ_* is the unconditioned state transformation induced by the measurement. The following result is a direct consequence of the results of, e.g.,¹⁶:

Theorem 2. *Let (Ω_i, Σ_i) , $i = 1, 2$, and (Ω, Σ) be measurable spaces and \mathcal{H} and \mathcal{K} separable Hilbert spaces.*

(i) Suppose that $M_i : \Sigma_i \rightarrow \mathcal{L}(\mathcal{H})$, $i = 1, 2$, are jointly measurable observables in \mathcal{H} . Let (\mathcal{M}, P, J) be any minimal Naimark dilation for M_1 . Fix a joint observable N for M_1 and M_2 . There is a unique POVM $F : \Sigma_2 \rightarrow \mathcal{L}(\mathcal{M})$ such that $P(X)F(Y) = F(Y)P(X)$ for all $X \in \Sigma_1$ and $Y \in \Sigma_2$ and

$$N(X \times Y) = J^*P(X)F(Y)J, \quad X \in \Sigma_1, \quad Y \in \Sigma_2.$$

(ii) Suppose that an observable $M : \Sigma \rightarrow \mathcal{L}(\mathcal{H})$ and a channel $\Phi : \mathcal{L}(\mathcal{K}) \rightarrow \mathcal{L}(\mathcal{H})$ are compatible and fix a joint instrument $\mathcal{J} : \mathcal{L}(\mathcal{K}) \times \Sigma \rightarrow \mathcal{L}(\mathcal{H})$ for M and Φ , any minimal Naimark dilation (\mathcal{M}, P, J) for M , and any minimal Stinespring dilation (\mathcal{K}', V) for Φ . There is a unique channel $T : \mathcal{L}(\mathcal{K}) \rightarrow \mathcal{L}(\mathcal{M})$ and a unique POVM $E : \Sigma \rightarrow \mathcal{L}(\mathcal{K}')$ such that $T(B)P(X) = P(X)T(B)$ for all $B \in \mathcal{L}(\mathcal{K})$ and $X \in \Sigma$ and

$$\mathcal{J}(B, X) = J^*T(B)P(X)J = V^*(B \otimes E(X))V, \quad B \in \mathcal{L}(\mathcal{K}), \quad X \in \Sigma.$$

When we combine item (ii) of the above theorem with Theorem 1, we obtain the following result; see also³² (Theorem 1)

Theorem 3. Let (Ω, Σ) be a measurable space, \mathcal{H} a separable Hilbert space, and $M : \Sigma \rightarrow \mathcal{L}(\mathcal{H})$ an observable. Pick the minimal Naimark dilation $(\mathcal{H}_\oplus, P_\oplus, J_\oplus)$ of Theorem 1 for M . Let $\mathcal{J} : \mathcal{L}(\mathcal{K}) \times \Sigma \rightarrow \mathcal{L}(\mathcal{H})$ be an M -instrument. There is a unique channel $T : \mathcal{L}(\mathcal{K}) \rightarrow \mathcal{L}(\mathcal{H}_\oplus)$ defined by a (weakly μ -measurable) field $x \mapsto T_x$ of channels $T_x : \mathcal{L}(\mathcal{K}) \rightarrow \mathcal{L}(\mathcal{H}(x))$,

$$T(B) = \int_{\Omega}^{\oplus} T_x(B) d\mu(x),$$

i.e., $(T(B)\psi)(x) = T_x(B)\psi(x)$ for all $B \in \mathcal{L}(\mathcal{K})$, $\psi \in \mathcal{H}_\oplus$, and μ -a.a. $x \in \Omega$, such that

$$\mathcal{J}(B, X) = J_\oplus^*T(B)P_\oplus(X)J_\oplus, \quad B \in \mathcal{L}(\mathcal{K}), \quad X \in \Sigma.$$

A. Connection between joint and sequential measurements

A special case of joint observables is sequential measurements where an initial observable $M : \Sigma \rightarrow \mathcal{L}(\mathcal{H})$ is first measured yielding some M -instrument $\mathcal{J} : \mathcal{L}(\mathcal{K}) \times \Sigma \rightarrow \mathcal{L}(\mathcal{H})$ with output Hilbert space \mathcal{K} . Then some observable $M' : \Sigma' \rightarrow \mathcal{L}(\mathcal{K})$ is measured. The conditional probability for obtaining an outcome within $Y \in \Sigma'$ in the second measurement, conditioned by the first measurement observing a value in $X \in \Sigma$, is

$$\text{tr}[\mathcal{J}_*(\rho, X)M'(Y)] = \text{tr}[\rho\mathcal{J}(M'(Y), X)]$$

when the system is initially in the state $\rho \in \mathcal{S}(\mathcal{H})$. For all spaces (Ω, Σ) and (Ω', Σ') , the positive operator bimeasure $(X, Y) \mapsto \mathcal{J}(\mathbf{M}'(Y), X)$ extends into a POVM on $\Sigma \otimes \Sigma'^{28,38}$. In this case, the extension $\mathbf{J} : \Sigma \otimes \Sigma' \rightarrow \mathcal{L}(\mathcal{H})$ is a joint observable for the initial observable \mathbf{M} and a distorted version $\mathbf{M}'' = \mathcal{J}(\mathbf{M}'(\cdot), \Omega)$ of the second observable. As shown in Section I, any joint measurement of discrete observables can be implemented as a sequential measurement; see also¹⁸ for this fact and its generalizations in the case of discrete observables. Next we show that joint and sequential measurements are, in this sense, equivalent in a very general case. Whenever (Ω, Σ) is a measurable space and μ is a probability measure on Σ , we denote by \mathbf{P}_μ the canonical spectral measure on $L^2(\mu)$, i.e., $(\mathbf{P}_\mu(X)\psi)(x) = \chi_X(x)\psi(x)$ for all $X \in \Sigma$, $\psi \in L^2(\mu)$, and μ -a.a. $x \in \Omega$. When \mathcal{K} is a Hilbert space we naturally identify $L^2(\mu) \otimes \mathcal{K}$ with the L^2 -space of functions $\Omega \rightarrow \mathcal{K}$.

Proposition 2. *Suppose that (Ω_i, Σ_i) , $i = 1, 2$, are countably generated measurable spaces and \mathcal{H} is a separable Hilbert space. Assume that $\mathbf{M}_i : \Sigma_i \rightarrow \mathcal{L}(\mathcal{H})$, $i = 1, 2$, are jointly measurable observables with a joint observable \mathbf{N} . There is a separable Hilbert space \mathcal{K} , an \mathbf{M}_1 -instrument $\mathcal{J} : \mathcal{L}(\mathcal{K}) \times \Sigma_1 \rightarrow \mathcal{L}(\mathcal{H})$, and a POVM $\mathbf{M}' : \Sigma_2 \rightarrow \mathcal{L}(\mathcal{K})$ such that*

$$\mathbf{N}(X \times Y) = \mathcal{J}(\mathbf{M}'(Y), X), \quad X \in \Sigma_1, \quad Y \in \Sigma_2. \quad (1)$$

Proof. Choose probability measures $\mu_i : \Sigma \rightarrow [0, 1]$ such that $\mathbf{M}_i \ll \mu_i$, $i = 1, 2$. Pick a minimal Naimark dilation $(\mathcal{H}_\oplus, \mathbf{P}_\oplus, J_\oplus)$ of Theorem 1 for \mathbf{M}_1 where

$$\mathcal{H}_\oplus := \int_{\Omega_1}^{\oplus} \mathcal{H}(x) d\mu_1(x)$$

is a direct integral space which is separable since (Ω_1, Σ_1) is countably generated. According to Theorem 2, there is a POVM $\mathbf{F} : \Sigma_2 \rightarrow \mathcal{L}(\mathcal{H}_\oplus)$ such that $\mathbf{P}_\oplus(X)\mathbf{F}(Y) = \mathbf{F}(Y)\mathbf{P}_\oplus(X)$ and $\mathbf{N}(X \times Y) = J_\oplus^* \mathbf{P}_\oplus(X)\mathbf{F}(Y)J_\oplus$ for all $X \in \Sigma_1$ and $Y \in \Sigma_2$. Let $(\mathcal{M}, \mathbf{Q}, K)$ be a minimal Naimark dilation for \mathbf{F} . Again, \mathcal{M} is separable since \mathcal{H}_\oplus is separable and Σ_2 is countably generated.

Fix $X \in \Sigma_1$ and define $\mathbf{F}_X : \Sigma_2 \rightarrow \mathcal{L}(\mathcal{H}_\oplus)$ by $\mathbf{F}_X(Y) = \mathbf{P}_\oplus(X)\mathbf{F}(Y)$. Now $\mathbf{F}_X(Y) \leq \mathbf{F}(Y)$ for all $Y \in \Sigma_2$, so that one can define a unique $\tilde{\mathbf{P}}(X) \in \mathcal{L}(\mathcal{M})$ by $\tilde{\mathbf{P}}(X)\mathbf{Q}(Y)K\psi := \mathbf{Q}(Y)K\mathbf{P}_\oplus(X)\psi$, $Y \in \Sigma_2$ and $\psi \in \mathcal{H}_\oplus$ (see, e.g., a similar proof of¹⁶ (Proposition 2.1)). Clearly, $\tilde{\mathbf{P}}(X)^2 = \tilde{\mathbf{P}}(X)$, $\tilde{\mathbf{P}}(X)\mathbf{Q}(Y) = \mathbf{Q}(Y)\tilde{\mathbf{P}}(X)$, and $\mathbf{F}_X(Y) = K^*\tilde{\mathbf{P}}(X)\mathbf{Q}(Y)K$ for all $Y \in \Sigma_2$. Hence, $X \mapsto \tilde{\mathbf{P}}(X)$ is a PVM.

For all $X \in \Sigma_1$ and $Y \in \Sigma_2$, define the projection $R(X, Y) = \tilde{P}(X)Q(Y) \in \mathcal{L}(\mathcal{M})$. Since \mathbf{N} is a POVM, for any $\varphi_1, \varphi_2 \in \mathcal{H}$, $X_1, X_2 \in \Sigma_1$, and $Y_1, Y_2 \in \Sigma_2$, the complex bimeasure

$$\begin{aligned} (X, Y) &\mapsto \langle Q(Y_1)KP_{\oplus}(X_1)J_{\oplus}\varphi_1 | R(X, Y)Q(Y_2)KP_{\oplus}(X_2)J_{\oplus}\varphi_2 \rangle \\ &= \langle \varphi_1 | \mathbf{N}((X \times Y) \cap (X_1 \times Y_1) \cap (X_2 \times Y_2)) \varphi_2 \rangle. \end{aligned}$$

extends into a complex measure on $\Sigma_1 \otimes \Sigma_2$. Using the minimality of the subsequent dilations, one finds that $(X, Y) \mapsto \langle \xi | R(X, Y) \xi \rangle$ extends into a measure for all $\xi \in \mathcal{M}$. Thus $(X, Y) \mapsto \tilde{P}(X)Q(Y)$ extends into a PVM which we shall also denote by R .

Since \mathcal{M} is separable, we may diagonalize R and thus identify \mathcal{M} with the direct integral space

$$\mathcal{M}_{\oplus} = \int_{\Omega_1 \times \Omega_2}^{\oplus} \mathcal{M}(x, y) d(\mu_1 \times \mu_2)(x, y),$$

where R operates as the canonical spectral measure. From now on, let us fix a separable infinite-dimensional Hilbert space \mathcal{M}_{∞} so that we may define a decomposable isometry $W : \mathcal{M} \rightarrow \overline{\mathcal{M}} := L^2(\mu_1 \times \mu_2) \otimes \mathcal{M}_{\infty} \cong L^2(\mu_1) \otimes [L^2(\mu_2) \otimes \mathcal{M}_{\infty}]$,

$$W = \int_{\Omega_1 \times \Omega_2}^{\oplus} W(x, y) d(\mu_1 \times \mu_2)(x, y),$$

where $W(x, y) : \mathcal{M}(x, y) \rightarrow \mathcal{M}_{\infty}$ are isometries. One may also define the decomposable isometry $W_1 : \mathcal{H}_{\oplus} \rightarrow \overline{\mathcal{M}}$, $W_1 = \int_{\Omega_1}^{\oplus} W_1(x) d\mu_1(x)$, where $W_1(x) : \mathcal{H}(x) \rightarrow L^2(\mu_2) \otimes \mathcal{M}_{\infty}$ are isometries, and $\overline{K} := WKW_1^* \in \mathcal{L}(\overline{\mathcal{M}})$.

Define the canonical spectral measure $\overline{R} := P_{\mu_1 \times \mu_2} \otimes I_{\mathcal{M}_{\infty}}$ of $\overline{\mathcal{M}}$ with the margin $\overline{P} : \Sigma_1 \rightarrow \mathcal{L}(\overline{\mathcal{M}})$, $\overline{P}(X) = \overline{R}(X \times \Omega_2) = P_{\mu_1}(X) \otimes I_{L^2(\mu_2)} \otimes I_{\mathcal{M}_{\infty}}$. It is simple to check that $\overline{R}(Z)W = WR(Z)$ and $P_{\oplus}(X)W_1^* = W_1^*\overline{P}(X)$ for all $Z \in \Sigma_1 \otimes \Sigma_2$ and $X \in \Sigma_1$. This means that

$$\overline{P}(X)\overline{K} = W\tilde{P}(X)KW_1^* = WKP_{\oplus}(X)W_1^* = \overline{K}\overline{P}(X)$$

for all $X \in \Sigma_1$. Thus, $\overline{K} = \int_{\Omega_1}^{\oplus} \overline{K}(x) d\mu_1(x)$ where $\overline{K}(x) \in \mathcal{L}(L^2(\mu_2) \otimes \mathcal{M}_{\infty})$. Define the isometry $\tilde{K} := WK = \overline{K}W_1 = \int_{\Omega_1}^{\oplus} \tilde{K}(x) d\mu(x)$ with the isometries $\tilde{K}(x) = \overline{K}(x)W_1(x) : \mathcal{H}(x) \rightarrow L^2(\mu_2) \otimes \mathcal{M}_{\infty}$.

For μ_1 -a.a. $x \in \Omega_1$ define the channel

$$T_x : \mathcal{L}(L^2(\mu_2)) \rightarrow \mathcal{L}(\mathcal{H}(x)), \quad B \mapsto T_x(B) := \tilde{K}(x)^*(B \otimes I_{\mathcal{M}_{\infty}})\tilde{K}(x).$$

Since the field $x \mapsto \tilde{K}(x)$ of isometries is measurable, one may define the channel

$$T : \mathcal{L}(L^2(\mu_2)) \rightarrow \mathcal{L}(\mathcal{H}_{\oplus}), \quad B \mapsto T(B) := \int_{\Omega}^{\oplus} T_x(B) d\mu(x).$$

Using the intertwining properties of the various isometries and POVMs we have, for all $\varphi \in \mathcal{H}$, $X \in \Sigma_1$, and $Y \in \Sigma_2$,

$$\begin{aligned}
\langle J_{\oplus}\varphi | \mathbf{P}_{\oplus}(X)T(\mathbf{P}_{\mu_2}(Y))J_{\oplus}\varphi \rangle &= \int_X \langle (J_{\oplus}\varphi)(x) | T_x(\mathbf{P}_{\mu_2}(Y))(J_{\oplus}\varphi)(x) \rangle d\mu_1(x) \\
&= \int_X \langle \tilde{K}(x)(J_{\oplus}\varphi)(x) | (\mathbf{P}_{\mu_2}(Y) \otimes I_{\mathcal{M}_{\infty}})\tilde{K}(x)(J_{\oplus}\varphi)(x) \rangle d\mu_1(x) \\
&= \langle \tilde{K}J_{\oplus}\varphi | \bar{\mathbf{R}}(X \times Y)\tilde{K}J_{\oplus}\varphi \rangle = \langle KJ_{\oplus}\varphi | \mathbf{R}(X \times Y)KJ_{\oplus}\varphi \rangle \\
&= \langle J_{\oplus}\varphi | \mathbf{P}_{\oplus}(X)\mathbf{F}(Y)J_{\oplus}\varphi \rangle = \langle \varphi | \mathbf{N}(X \times Y)\varphi \rangle.
\end{aligned}$$

Hence, $\mathbf{N}(X \times Y) = J_{\oplus}^*\mathbf{P}_{\oplus}(X)T(\mathbf{P}_{\mu_2}(Y))J_{\oplus}$ for all $X \in \Sigma_1$ and $Y \in \Sigma_2$.

Define the instrument $\mathcal{J} : \mathcal{L}(L^2(\mu_2)) \times \Sigma_1 \rightarrow \mathcal{L}(\mathcal{H})$ by $\mathcal{J}(B, X) = J_{\oplus}^*\mathbf{P}_{\oplus}(X)T(B)J_{\oplus}$, see³² (Theorem 1). The choices $\mathcal{K} := L^2(\mu_2)$ and $M' := \mathbf{P}_{\mu_2}$ yield Equation (1). \square

IV. OBSERVABLES DETERMINING THE FUTURE

We now turn our attention to those observables which have the property that, no matter how we measure them, registering an outcome unequivocally determines the post-measurement state of the system under study. Let us first introduce a special class of observables.

Definition 3. Let an $\mathbf{M} \in \text{Obs}(\Sigma, \mathcal{H})$ be associated with the Naimark dilation $(\mathcal{H}_{\oplus}, \mathbf{P}_{\oplus}, J_{\oplus})$ of Theorem 1. If $\dim \mathcal{H}(x) = 1$ for μ -a.a. $x \in \Omega$, we say that \mathbf{M} is of *rank 1*. In this case, $\mathcal{H}_{\oplus} = L^2(\mu)$ and $\mathbf{P}_{\oplus} = \mathbf{P}_{\mu}$.

Let the observable \mathbf{M} of Theorem 1 be of rank 1. Also assume that $\mathcal{J} : \mathcal{L}(\mathcal{K}) \times \Sigma \rightarrow \mathcal{L}(\mathcal{H})$ is an \mathbf{M} -instrument defined by the pointwise channels $T_x : \mathcal{L}(\mathcal{K}) \rightarrow \mathcal{L}(\mathcal{H}(x))$ of Theorem 3. Because of the rank-1 assumption, there are states $\sigma_x \in \mathcal{S}(\mathcal{K})$ such that $T_x(B) = \text{tr}[\sigma_x B]$, $x \in \Omega$, $B \in \mathcal{L}(\mathcal{K})$. It follows that \mathcal{J} is of the following type:

Definition 4. Let \mathcal{H} and \mathcal{K} be Hilbert spaces and (Ω, Σ) a measurable space. We say that an instrument $\mathcal{J} : \mathcal{L}(\mathcal{K}) \times \Sigma \rightarrow \mathcal{L}(\mathcal{H})$ is *nuclear* if there is a weakly μ -measurable field $\Omega \ni x \mapsto \sigma_x \in \mathcal{S}(\mathcal{K})$ of states (meaning that all maps $x \mapsto \text{tr}[\sigma_x B]$, $B \in \mathcal{L}(\mathcal{K})$, are μ -measurable) such that

$$\mathcal{J}(B, X) = \int_X \text{tr}[\sigma_x B] d\mathbf{M}(x), \quad X \in \Sigma, \quad B \in \mathcal{L}(\mathcal{K}).$$

The term *nuclear* follows the terminology of Cycon and Hellwig⁸. The above definition means that, in the Schrödinger picture, a nuclear instrument \mathcal{J} has the form

$$\mathcal{J}_*(\rho, X) = \int_X \sigma_x \mathrm{d}p_\rho^{\mathbf{M}}(x), \quad \rho \in \mathcal{S}(\mathcal{H}), \quad X \in \Sigma,$$

where the integral is defined weakly. Physically this means that a nuclear instrument prepares the quantum system into some post-measurement state which solely depends on the outcome registered, not on the pre-measurement state of the system. This is why also the name *measure-and-prepare instrument* could also be used. Thus, any measurement of a rank-1 observable is described by a nuclear instrument and registering a value fully determines the post-measurement state. This is to say, rank-1 observables determine the future of the system under measurement. In fact, also the contrary is true as the following result from³² tells us.

Theorem 4. *An observable $\mathbf{M} : \Sigma \rightarrow \mathcal{L}(\mathcal{H})$ is rank-1 if and only if each \mathbf{M} -instrument $\mathcal{J} : \mathcal{L}(\mathcal{K}) \times \Sigma \rightarrow \mathcal{L}(\mathcal{H})$ is nuclear (where \mathcal{K} is any Hilbert space).*

The above result can be reformulated in the form that *an observable determines the future if and only if it is of rank 1*. The channel $\mathcal{J}(\cdot, \Omega)$ associated with a measurement of a rank-1 observable is also seen to be entanglement breaking²².

Let $\mathbf{M} \in \text{Obs}(\Sigma, \mathcal{H})$ with vectors $d_k(x)$ of Theorem 1, $\Omega^1 := \mathbb{N} \times \Omega$, and let Σ^1 be the product σ -algebra of $2^{\mathbb{N}}$ and Σ . Let $\mu^1 : \Sigma^1 \rightarrow [0, \infty]$ be the product measure of the counting measure and μ . Define $d(k, x) = d_k(x)$ if $k < m(x) + 1$ and $d(k, x) = 0$ if $k > m(x)$. Then

$$\langle \varphi | \mathbf{M}^1(X^1) \psi \rangle = \int_{X^1} \langle \varphi | d(k, x) \rangle \langle d(k, x) | \psi \rangle \mathrm{d}\mu^1(k, x), \quad \varphi, \psi \in V_{\mathbf{h}}, \quad X^1 \in \Sigma^1, \quad (2)$$

defines a rank-1 POVM $\mathbf{M}^1 : \Sigma^1 \rightarrow \mathcal{L}(\mathcal{H})$; we say that \mathbf{M}^1 is a *maximally refined version* of \mathbf{M} .

Since $\mathbf{M}(X) = \mathbf{M}^1(f^{-1}(X))$ where $f : \Omega^1 \rightarrow \Omega$ is a measurable function defined by $f(k, x) = f(x)$ for all $k \in \mathbb{N}$ and $x \in \Omega$, \mathbf{M} is a relabeling of \mathbf{M}^1 . Note that the value space of \mathbf{M}^1 contains the multiplicities (k, x) , $k < m(x) + 1$, of a measurement outcome x of \mathbf{M} . Moreover, \mathbf{M} and \mathbf{M}^1 are jointly measurable and \mathbf{M}^1 can be measured by performing a sequential measurement of \mathbf{M} and some discrete ‘multiplicity’ observable³⁴. We will see that the maximally refined version of an observable possesses many of the same optimality properties as the original observable meaning that we may freely assume the rank-1 property for these observables.

V. POST-PROCESSING AND POST-PROCESSING MAXIMALITY

Let us begin with a definition.

Definition 5. Let (Ω_1, Σ_1) and (Ω_2, Σ_2) be measurable spaces. Also assume that $\mu : \Sigma_1 \rightarrow \mathbb{R}$ is a positive measure. We say that a map $\beta : \Sigma_2 \times \Omega_1 \rightarrow \mathbb{R}$ is a μ -weak Markov kernel²⁴ if

- (i) $\beta(Y, \cdot) : \Omega_1 \rightarrow \mathbb{R}$ is μ -measurable for all $Y \in \Sigma_2$,
- (ii) $\beta(Y, x) \geq 0$ for all $Y \in \Sigma_2$ and μ -a.a. $x \in \Omega_1$,
- (iii) $\beta(\Omega_2, x) = 1$ for μ -a.a. $x \in \Omega_1$, and
- (iv) for all pairwise disjoint sequences $Y_1, Y_2, \dots \in \Sigma_2$,

$$\beta\left(\bigcup_{j=1}^{\infty} Y_j, x\right) = \sum_{j=1}^{\infty} \beta(Y_j, x)$$

for μ -a.a. $x \in \Omega_1$.

If $\beta(\cdot, x)$ is a probability measure for all $x \in \Omega_1$ and the maps $\beta(Y, \cdot)$ are measurable then β is simply called a *Markov kernel*.

When μ_1 is a probability measure on (Ω_1, Σ_1) , $\mu_1 \ll \mu$, and $\beta : \Sigma_2 \times \Omega_1 \rightarrow \mathbb{R}$ is a μ -weak Markov kernel, then the set function

$$\Sigma_1 \times \Sigma_2 \ni (X, Y) \mapsto B(X, Y) := \int_X \beta(Y, x) d\mu_1(x) \in [0, 1]$$

is a *probability bimeasure* with the marginal probability measures $X \mapsto B(X, \Omega_2) = \mu_1(X)$ and $Y \mapsto B(\Omega_1, Y) =: \mu_1^\beta(Y)$. Recall that $B : \Sigma_1 \times \Sigma_2 \rightarrow \mathbb{C}$ is a bimeasure if $B(X, \cdot)$, $X \in \Sigma_1$, and $B(\cdot, Y)$, $Y \in \Sigma_2$, are (complex) measures. As an immediate consequence of Carathéodory's extension theorem, one gets the well-know result stating that if β is a Markov kernel then B extends into probability measure $\bar{B} : \Sigma_1 \times \Sigma_2 \rightarrow [0, 1]$, i.e., $\bar{B}(X \times Y) = B(X, Y)$ for all $X \in \Sigma_1$ and $Y \in \Sigma_2$. Note that μ_1^β can be interpreted as a result of (classical) data processing represented by β . We call this data processing scene *post-processing* since the processing can be carried out after obtaining the data represented by the measure μ_1 . This data processing scheme generalizes to the case of POVMs in the following way.

Definition 6. Let $M_1 : \Sigma_1 \rightarrow \mathcal{L}(\mathcal{H})$ be an observable operating in the Hilbert space \mathcal{H} . We assume that there is a (probability) measure μ on (Ω, Σ) such that $M_1 \ll \mu$. We say that an observable $M_2 : \Sigma_2 \rightarrow \mathcal{L}(\mathcal{H})$ is a *post-processing of M_1* , if there is a μ -weak Markov kernel $\beta : \Sigma_2 \times \Omega_1 \rightarrow \mathbb{R}$ such that $p_\rho^{M_2} = (p_\rho^{M_1})^\beta$ for all $\rho \in \mathcal{S}(\mathcal{H})$ or, equivalently,

$$M_2(Y) = \int_{\Omega_1} \beta(Y, x) dM_1(x) \quad (\text{weakly})$$

for all $Y \in \Sigma_2$. We denote $M_2 = M_1^\beta$.

The above means that by measuring M_1 , we obtain all the information obtainable by measuring M_2 ; we just have to process the data given by M_1 classically with the fixed kernel β . Thus, M_1 can give us at least the same amount of information on the quantum system as M_2 modulo classical data processing. Note that if M_2 is a relabeling of M_1 , i.e. $M_2(Y) = M_1(f^{-1}(Y))$, then $M_2 = M_1^\beta$ where $\beta(Y, x) = \chi_{f^{-1}(Y)}(x)$ is a Markov kernel.

We may thus set up an information-content ‘order’ among observables^{3,13,29} $M_2 \leq_{\text{post}} M_1$ if there is a μ -weak Markov kernel $\beta : \Sigma_2 \times \Omega_1 \rightarrow \mathbb{R}$ (where $M_1 \ll \mu$) such that $M_2 = M_1^\beta$. We may also say that M_1 and M_2 are *post-processing equivalent* if there are weak Markov kernels β and γ such that $M_2 = M_1^\beta$ and $M_1 = M_2^\gamma$. Recall that the ‘order’ \leq_{post} here may not actually be a partial order (because of the failure of transitivity); for situations where this problem can be overcome and identification of canonical representatives of the resulting equivalence classes, see²⁷. An observable M is *post-processing maximal* or *post-processing clean* if, for any observable M' such that $M \leq_{\text{post}} M'$, one has $M' \leq_{\text{post}} M$. The maximal observables have been characterized earlier in the case of discrete outcomes¹³ (Theorem 3.4). We generalize this characterization for observables with nice outcome spaces. For that, we need the following results.

Proposition 3. *Let (Ω_1, Σ_1) be nice, (Ω_2, Σ_2) countably generated, and $B : \Sigma_1 \times \Sigma_2 \rightarrow [0, 1]$ a probability bimeasure. Denote $\mu_1 = B(\cdot, \Omega_2)$.*

(i) *There exists a probability measure $\bar{B} : \Sigma_1 \otimes \Sigma_2 \rightarrow [0, 1]$ such that $\bar{B}(X \times Y) = B(X, Y)$ for all $X \in \Sigma_1$ and $Y \in \Sigma_2$.*

(ii) *There exists a Markov kernel $\beta : \Sigma_2 \times \Omega_1 \rightarrow [0, 1]$ such that $B(X, Y) = \int_X \beta(Y, x) d\mu_1(x)$ for all $X \in \Sigma_1$ and $Y \in \Sigma_2$.*

Proof. First we note that (i) holds in the case where (Ω_1, Σ_1) and (Ω_2, Σ_2) are standard Borel spaces¹⁰ (Lemma 4.2.1) showing that Lemma 12.1 of³⁷ holds even in the case where probability measures on $\Sigma_1 \otimes \Sigma_2$ (i.e. joint probability measures) are replaced with probability bimeasures on $\Sigma_1 \times \Sigma_2$. Manifestly the rest of the proof of Theorem 12.1 of³⁷ can be carried out by replacing joint probability measures with probability bimeasures everywhere. This proves item (ii). Item (i) follows from (ii) by recalling the well-known fact that any Markov kernel defines a joint probability measure. \square

Remark 2. Let (Ω_1, Σ_1) and (Ω_2, Σ_2) be as in Proposition 3, $M_i \in \text{Obs}(\Sigma_i, \mathcal{H})$, $i = 1, 2$, $M_1 \sim \mu_1$, and $M_2 = M_1^\beta$ where β is a μ_1 -weak Markov kernel, i.e. M_2 is a post-processing of M_1 . Since β defines a probability bimeasure, we immediately get from Proposition 3 the following results:

- There is a Markov kernel β' such that $M_2 = M_1^{\beta'}$ and $\beta(Y, x) = \beta'(Y, x)$ for all $Y \in \Sigma_2$ and μ_1 -a.a. $x \in \Omega_1$.
- The POVMs M_1 and M_2 are jointly measurable, a joint observable $N \in \text{Obs}(\Sigma_1 \otimes \Sigma_2, \mathcal{H})$ being defined through $N(X \times Y) := \int_X \beta(Y, x) dM_1(x)$. Note also that the joint POVM N is a post-processing of M_1 , i.e. $N = M_1^\gamma$ where the (weak) Markov kernel γ is defined by $\gamma(X \times Y, x) := \chi_X(x)\beta(Y, x)$, $X \in \Sigma_1$, $Y \in \Sigma_2$, $x \in \Omega_1$.

For the remainder of this section, let us fix measurable spaces (Ω_i, Σ_i) , $i = 1, 2$, a probability measure $\mu_1 : \Sigma_1 \rightarrow [0, 1]$ and a Markov kernel $\beta : \Sigma_2 \times \Omega_1 \rightarrow [0, 1]$. Let us denote $\mu_2 := \mu_1^\beta$ and by $\mu : \Sigma_1 \otimes \Sigma_2 \rightarrow [0, 1]$ the probability measure extended from the probability bimeasure

$$\Sigma_1 \times \Sigma_2 \ni (X, Y) \mapsto \int_X \beta(Y, x) d\mu_1(x) \in [0, 1].$$

Let $Z \in \Sigma_1 \otimes \Sigma_2$ and, for each $x \in \Omega_1$, denote $Z_x := \{y \in \Omega_2 \mid (x, y) \in Z\}$. It is immediate that $Z_x \in \Sigma_2$ for each $x \in \Omega_1$. Moreover, the function $\Omega_1 \ni x \mapsto \beta(Z_x, x) \in [0, 1]$ is measurable for all $Z \in$

$\text{Sigma}_1 \otimes \text{Sigma}_2$. This is easy to show by first showing this for $Z = X \times Y$, $X \in \Sigma_1$, $Y \in \Sigma_2$, and then proving that the family of sets $Z \in \Sigma_1 \otimes \Sigma_2$ such that the above function is measurable is a Dynkin class (a collection of sets containing $\Omega_1 \times \Omega_2$ and closed under complementation and countable disjoint unions) and then using the Dynkin class theorem.

We may now make an important observation: for any $Z \in \Sigma_1 \otimes \Sigma_2$,

$$\mu(Z) = \int_{\Omega_1} \beta(Z_x, x) d\mu_1(x). \quad (3)$$

From these considerations using standard techniques, one may prove the following (well-known) generalized Tonelli theorem:

Theorem 5. *Whenever $h : \Omega_1 \times \Omega_2 \rightarrow [0, \infty]$ is measurable,*

$$\int_{\Omega_1 \times \Omega_2} h d\mu = \int_{\Omega_1} \int_{\Omega_2} h(x, y) \beta(dy, x) d\mu_1(x).$$

The following simple lemma is needed in Section V B.

Lemma 1. *Let $f \in L^\infty(\mu_1)$ and $g \in L^\infty(\mu_2)$. Define the complex measure $\nu_f : \Sigma_2 \rightarrow [0, 1]$,*

$$\nu_f(Y) = \int_{\Omega_1} \beta(Y, x) f(x) d\mu_1(x), \quad Y \in \Sigma_2.$$

One has

$$\int_{\Omega_1 \times \Omega_2} f(x)g(y) d\mu(x, y) = \int_{\Omega_2} g(y) d\nu_f(y).$$

Proof. The procedure is standard: First consider the case where f and g are simple functions. This first step is easily proven. Let then $f \geq 0$, $g \geq 0$, and let $f_n \geq 0$ and $g_n \geq 0$, $n \in \mathbb{N}$, be simple functions such that $f_n \uparrow f$ and $g_n \uparrow g$ pointwisely and uniformly respectively in μ_1 and in μ_2 . Fix $\varepsilon > 0$ and let $n_0 \in \mathbb{N}$ be such that (*grâce à* monotone convergence theorem)

$$\int_{\Omega_1 \times \Omega_2} f(x)g(y) d\mu(x, y) - \int_{\Omega_1 \times \Omega_2} f_n(x)g_n(y) d\mu(x, y) < \varepsilon/3,$$

$f(x) - f_n(x) < (3\|g\|_\infty)^{-1}\varepsilon$ for μ_1 -a.a. $x \in \Omega_1$, and $g(y) - g_n(y) < (3\|f\|_\infty)^{-1}\varepsilon$ for μ_2 -a.a. $y \in \Omega_2$ (hence also ν_f -a.e., as one easily checks) for every $n \geq n_0$. We now have for every $n \geq n_0$

$$\begin{aligned} & \left| \int_{\Omega_1 \times \Omega_2} f(x)g(y) d\mu(x, y) - \int_{\Omega_2} g d\nu_f \right| \\ & \leq \left| \int_{\Omega_1 \times \Omega_2} f(x)g(y) d\mu(x, y) - \int_{\Omega_1 \times \Omega_2} f_n(x)g_n(y) d\mu(x, y) \right| \\ & + \left| \int_{\Omega_2} g d\nu_f - \int_{\Omega_1 \times \Omega_2} f_n(x)g_n(y) d\mu(x, y) \right| < \varepsilon/3 + \int_{\Omega_2} (g - g_n) d\nu_f + \int_{\Omega_2} g_n d\nu_{f-f_n} \leq \varepsilon \end{aligned}$$

using the fact that $\nu_f \leq \|f\|_\infty \mu_2$. This proves the second stage. The extension to complex-valued functions is straight-forward. \square

A. Joint measurements of rank-1 observables

For the results of the rest of this section, it is useful, as an interlude, to now turn our attention to joint-measurability issues of rank-1 observables. Let $M_i : \Sigma_i \rightarrow \mathcal{L}(\mathcal{H})$, $i = 1, 2$, be jointly measurable observables where M_1 is of rank 1. Let \mathcal{H} be separable and $(\mathcal{H}_\oplus, P_\oplus, J_\oplus)$ be the minimal (diagonal) Naimark dilation of M_1 introduced in Theorem 1 with the vector field $x \mapsto d_1(x) =: d(x)$ so that, for all $X \in \Sigma_1$,

$$\langle \varphi | M_1(X) \psi \rangle = \int_X \langle \varphi | d(x) \rangle \langle d(x) | \psi \rangle d\mu_1(x) = \int_X \overline{(J_\oplus \varphi)(x)} (J_\oplus \psi)(x) d\mu_1(x)$$

where $\varphi, \psi \in V_h$, since $\mathcal{H}(x) \equiv \mathbb{C}$ implies $\mathcal{H}_\oplus = L^2(\mu_1)$ and $P_\oplus = P_{\mu_1}$. According to Theorem 2, there is a unique POVM $F : \Sigma_2 \rightarrow \mathcal{L}(L^2(\mu_1))$ such that $P_{\mu_1}(X)F(Y) = F(Y)P_{\mu_1}(X)$ for all $X \in \Sigma_1$ and $Y \in \Sigma_2$ and $M_2(Y) = J_\oplus^* F(Y) J_\oplus$ for all $Y \in \Sigma_2$. Hence, for any $Y \in \Sigma_2$, there is a measurable function $\beta(Y, \cdot) : \Omega_1 \rightarrow \mathbb{R}$ such that $(F(Y)\eta)(x) = \beta(Y, x)\eta(x)$ for all $\eta \in L^2(\mu_1)$ and μ_1 -a.a. $x \in \Omega_1$. It is simple to check that the map $\beta : \Sigma_2 \times \Omega_1 \rightarrow \mathbb{R}$ satisfies the conditions (i)–(iv) of Definition 5 implying that β is a μ_1 -weak Markov kernel and $M_2 = M_1^\beta$. Thus, we have³⁵:

Theorem 6. *Let $M : \Sigma \rightarrow \mathcal{L}(\mathcal{H})$ be a rank-1 observable of a separable \mathcal{H} . Any observable $M' : \Sigma' \rightarrow \mathcal{L}(\mathcal{H})$ jointly measurable with M is a post-processing of M .*

B. Post-processing clean observables

The general form of post-processing clean observables is claimed to have been solved in². There are, however, some problems in the definition of post-processing the paper uses: Despite the author's definition of post-processing involves, according to the terminology used here, weak Markov kernels, a kernel β is treated assuming that $\beta(\cdot, x)$ is a measure for a.a. x . Moreover, we find the proofs of the main theorems dubious. That is why we provide a new proof. We end up with the same characterization as in² though and use much of the same machinery as in². The next theorem is an essential part of the characterization of post-processing clean observables given in Corollary 2.

Theorem 7. *Let (Ω_i, Σ_i) , $i = 1, 2$, be measurable spaces, \mathcal{H} a separable Hilbert space, $M_1 \in \text{Obs}(\Sigma_1, \mathcal{H})$, and $\beta : \Sigma_2 \times \Omega_1 \rightarrow [0, 1]$ a Markov kernel. If M_1^β is of rank 1 then M_1 is of rank 1.*

Proof. Assume that μ_1 is a probability measure on (Ω_1, Σ_1) such that $\mathbf{M}_1 \ll \mu_1$. Clearly, $\mathbf{M}_2 := \mathbf{M}_1^\beta \ll \mu_2 := \mu_1^\beta$ (i.e. $\mu_2(Y) = \int_{\Omega_1} \beta(Y, x) d\mu_1(x)$). For any Hilbert-Schmidt operator $R \in \mathcal{L}(\mathcal{H})$, $Z \in \Sigma_i$, $i = 1, 2$, by the Radon-Nikodým property of the trace class,

$$R^* \mathbf{M}_i(Z) R = \int_X \mathbf{m}_i(z) d\mu_i(z),$$

where $\mathbf{m}_i : \Omega_i \rightarrow \mathcal{L}(\mathcal{H})$ is a weakly μ_i -measurable positive trace-class-valued function (which depends on R), see e.g.²³. Requiring \mathbf{M}_2 to be rank-1 is equivalent with $\mathbf{m}_2(y)$ being at most rank-1 almost everywhere. Fix now a Hilbert-Schmidt operator R and let \mathbf{m}_i be the corresponding densities of $R^* \mathbf{M}_i(\cdot) R$ with respect to μ_i . Let $\{\psi_j\}_j \subset \mathcal{H}$ be an orthonormal system whose closed linear span is the support of $R^* R$. Now

$$R^* \mathbf{M}_2(Y) R = \int_Y \mathbf{m}_2(y) d\mu_2(y) = \int_{\Omega_1} \beta(Y, x) \mathbf{m}_1(x) d\mu_1(x) \quad (4)$$

for all $Y \in \Sigma_2$. Since β is a Markov kernel, the probability bimeasure $(X, Y) \mapsto \int_X \beta(Y, x) d\mu_1(x)$ extends into a probability measure $\mu : \Sigma_1 \otimes \Sigma_2 \rightarrow [0, 1]$.

Define the functions $f_{i,j} \in L^\infty(\mu_1)$, $f_{i,j}(x) = \langle \psi_i | \mathbf{m}_1(x) \psi_j \rangle$, $x \in \Omega_1$. From (4) it follows that, using the notation of Lemma 1, $\nu_{f_{i,j}} = \langle R \psi_i | \mathbf{M}_2(\cdot) R \psi_j \rangle$. Define, for all $y \in \Omega_2$, $P(y)$ as the one-dimensional projection onto the support of $\mathbf{m}_2(y)$ and $P(y)^\perp = I_{\mathcal{H}} - P(y)$. Using the generalized Tonelli theorem 5, the polarization identity, the monotone convergence theorem, and Lemma 1, we obtain for every $\varphi \in \mathcal{H}$

$$\begin{aligned} & \int_{\Omega_1 \times \Omega_2} \langle P(y)^\perp \varphi | \mathbf{m}_1(x) P(y)^\perp \varphi \rangle d\mu(x, y) \\ &= \sum_{i,j} \int_{\Omega_1 \times \Omega_2} \langle P(y)^\perp \varphi | \psi_i \rangle \langle \psi_i | \mathbf{m}_1(x) \psi_j \rangle \langle \psi_j | P(y)^\perp \varphi \rangle d\mu(x, y) \\ &= \sum_{i,j} \int_{\Omega_2} \langle P(y)^\perp \varphi | \psi_i \rangle \langle \psi_j | P(y)^\perp \varphi \rangle d\nu_{f_{i,j}}(y) \\ &= \sum_{i,j} \int_{\Omega_2} \langle P(y)^\perp \varphi | \psi_i \rangle \langle \psi_j | P(y)^\perp \varphi \rangle \langle \psi_i | \mathbf{m}_2(y) \psi_j \rangle d\mu_2(y) \\ &= \int_{\Omega_2} \langle P(y)^\perp \varphi | \mathbf{m}_2(y) P(y)^\perp \varphi \rangle d\mu_2(y) = 0. \end{aligned}$$

Denote by Z^φ the set of those $(x, y) \in \Omega_1 \times \Omega_2$ such that $\langle P(y)^\perp \varphi | \mathbf{m}_1(x) P(y)^\perp \varphi \rangle = 0$; clearly, $Z^\varphi \in \Sigma_1 \otimes \Sigma_2$. Suppose that $\{\varphi_n\}_n \subset \mathcal{H}$ is a countable dense subset and define $Z := \bigcap_n Z^{\varphi_n}$. Easily one sees that, as $\mu(Z^{\varphi_n}) = 1$ for all n and also $\mu(Z) = 1$. Thus it follows that $P(y)^\perp \mathbf{m}_2(x) P(y)^\perp = 0$ for all $(x, y) \in Z$. Recalling the discussion before Theorem 5,

$$1 = \mu(Z) = \int_{\Omega_1} \beta(Z_x, x) d\mu_1(x)$$

implying that, for μ_1 -a.a. $x \in \Omega_1$, $\beta(Z_x, x) = 1$, so that, for μ_1 -a.a. $x \in \Omega_1$, $Z_x \neq \emptyset$. This means that, for μ_1 -a.a. $x \in \Omega_1$, there is $y \in \Omega_2$ such that $P(y)^\perp \mathbf{m}_1(x) P(y)^\perp = 0$ implying $\mathbf{m}_1(x) = P(y) \mathbf{m}_1(x) P(y)$, i.e., $\mathbf{m}_1(x)$ is at most rank-1. Since this holds for any Hilbert-Schmidt operator R , we have that \mathbf{M}_1 is rank-1. \square

From Remark 2 and the theorem above we get:

Corollary 1. *Suppose that (Ω_1, Σ_1) (resp. (Ω_2, Σ_2)) is a nice (resp. countably generated) measurable space and \mathcal{H} is a separable Hilbert space. Let $\mathbf{M}_i : \Sigma_i \rightarrow \mathcal{L}(\mathcal{H})$, $i = 1, 2$, be observables such that \mathbf{M}_2 is of rank 1. If \mathbf{M}_2 is a post-processing of \mathbf{M}_1 then \mathbf{M}_1 is of rank 1 as well.*

The next corollary gives an exhaustive characterization of post-processing clean observables with a nice value space. Especially, we find that such an observable is post-processing maximal if and only if it determines the future of the system under study.

Corollary 2. *Let (Ω, Σ) be a measurable space, \mathcal{H} a separable Hilbert space, and $\mathbf{M} \in \text{Obs}(\Sigma, \mathcal{H})$. If \mathbf{M} is of rank-1 then it is post-processing clean. The converse holds when (Ω, Σ) is nice.*

Proof. Suppose first that \mathbf{M} is rank-1 and $\mu \sim \mathbf{M}$ is a probability measure. Hence, according to Theorem 1, \mathbf{M} has a minimal Naimark dilation $(L^2(\mu), \mathbf{P}_\mu, J_\oplus)$. If \mathbf{M} is a post-processing of an $\tilde{\mathbf{M}} \in \text{Obs}(\tilde{\Sigma}, \mathcal{H})$ on some measurable space $(\tilde{\Omega}, \tilde{\Sigma})$, i.e. $\mathbf{M} = \tilde{\mathbf{M}}^{\tilde{\beta}}$ where $\tilde{\beta}$ is a $\tilde{\mu}$ -weak Markov kernel and $\tilde{\mu} \sim \tilde{\mathbf{M}}$, one can define a positive operator bimeasure

$$(X, Y) \mapsto \int_Y \tilde{\beta}(X, y) d\tilde{\mathbf{M}}(y) = J_\oplus^* \mathbf{P}_\mu(X) \mathbf{F}(Y) J_\oplus$$

where now \mathbf{F} is of the form $(\mathbf{F}(Y)\eta)(x) = \beta(Y, x)\eta(x)$ for all $Y \in \tilde{\Sigma}$, $\eta \in L^2(\mu)$ and μ -a.a. $x \in \Omega$, and thus $\tilde{\mathbf{M}} = \mathbf{M}^\beta$, see Section V A for details.

Assume now that \mathbf{M} is post-processing clean. Let $\mathbf{M}^1 : \Sigma^1 \rightarrow \mathcal{L}(\mathcal{H})$ be the rank-1 refinement of \mathbf{M} defined in (2) from which \mathbf{M} can be post-processed. Since \mathbf{M} is clean, \mathbf{M}^1 is also a post-processing of \mathbf{M} . If (Ω, Σ) is nice then (Ω^1, Σ^1) is nice (thus countably generated) and Corollary 1 implies that \mathbf{M} is rank-1 as well (i.e., \mathbf{M} and \mathbf{M}^1 coincide). \square

VI. OBSERVABLES DETERMINING THE PAST

In this section, we concentrate on observables that define the past of the system under study, i.e., those observables whose measurement outcome statistics completely determine the state of the system prior to the measurement.

Let \mathcal{H} be a Hilbert space and (Ω, Σ) a measurable space. Let $\mathbf{M} \in \text{Obs}(\Sigma, \mathcal{H})$ and recall our earlier definition $p_\rho^{\mathbf{M}} = \text{tr}[\rho \mathbf{M}(\cdot)]$ for all $\rho \in \mathcal{S}(\mathcal{H})$. Note that the map $\rho \mapsto p_\rho^{\mathbf{M}}$ is an affine map which is continuous with respect to the trace norm on $\mathcal{S}(\mathcal{H})$ and the total variation norm of probability measures. If this map is an injection, the natural conclusion is that the observable \mathbf{M} can separate all states; with different states of the system, the outcome statistics will always differ. How one can actually determine the state of the system prior to the measurement is not discussed here; the reader is redirected to²⁶ for this issue.

This prompts the following definition: an observable $\mathbf{M} : \Sigma \rightarrow \mathcal{L}(\mathcal{H})$ is *informationally complete* if for $\rho, \sigma \in \mathcal{S}(\mathcal{H})$, $\rho \neq \sigma$ implies $p_\rho^{\mathbf{M}} \neq p_\sigma^{\mathbf{M}}$. This injectivity extends to the whole of $\mathcal{T}(\mathcal{H})$, and thus informational completeness of \mathbf{M} is equivalent with the following: for any $T \in \mathcal{T}(\mathcal{H})$, the condition $\text{tr}[T\mathbf{M}(X)] = 0$ for all $X \in \Sigma$ implies $T = 0$. From this we see that the range $\text{ran } \mathbf{M} = \{\mathbf{M}(X) \mid X \in \Sigma\}$ of \mathbf{M} has to be extensive enough to separate the trace class $\mathcal{T}(\mathcal{H})$. Indeed, \mathbf{M} is informationally complete if and only if the ultraweak closure of the linear hull of $\text{ran } \mathbf{M}$ (which coincides with the double commutant $(\text{ran } \mathbf{M})''$) is the whole of $\mathcal{L}(\mathcal{H})$ ⁵ (Proposition 18.1). We can make the following important immediate observations: *If \mathbf{M} is informationally complete, its rank-1 refinement \mathbf{M}^1 is informationally complete as well, and any joint measurement of an informationally complete observable with some observable is also informationally complete. More generally, if a post-processing of an observable is informationally complete, then the post-processed observable is also informationally complete.*

To further quantify the informational content of an $\mathbf{M} \in \text{Obs}(\Sigma, \mathcal{H})$ in a Hilbert space \mathcal{H} , let us define for each $\rho \in \mathcal{S}(\mathcal{H})$ the set $[\rho]^{\mathbf{M}} \subseteq \mathcal{S}(\mathcal{H})$ as the set of those states $\sigma \in \mathcal{S}(\mathcal{H})$ such that $p_\sigma^{\mathbf{M}} = p_\rho^{\mathbf{M}}$. It is evident that \mathbf{M} is informationally complete if and only if $[\rho]^{\mathbf{M}} = \{\rho\}$ for all $\rho \in \mathcal{S}(\mathcal{H})$. This definition can be generalized to the case of sets \mathcal{O} of observables (in the same Hilbert space \mathcal{H}):

$$[\rho]^{\mathcal{O}} := \{\sigma \in \mathcal{S}(\mathcal{H}) \mid p_\sigma^{\mathbf{M}} = p_\rho^{\mathbf{M}}, \forall \mathbf{M} \in \mathcal{O}\}$$

One can say that a set \mathcal{O} of observables is informationally complete if $[\rho]^{\mathcal{O}} = \{\rho\}$ for all $\rho \in \mathcal{S}(\mathcal{H})$.

An observable $M : \Sigma \rightarrow \mathcal{L}(\mathcal{H})$ is said to be *commutative* if $M(X)M(Y) = M(Y)M(X)$ for all $X, Y \in \Sigma$. Let $\mathcal{L} \subseteq \mathcal{L}(\mathcal{H})$ be a set of selfadjoint operators. We call the set of vectors $\varphi \in \mathcal{H}$ such that $L_1 \cdots L_n \varphi = L_{\pi(1)} \cdots L_{\pi(n)} \varphi$ for any $L_1, \dots, L_n \in \mathcal{L}$, any permutation π of $\{1, \dots, n\}$, and any $n \in \mathbb{N}$ as the *commutation domain* of \mathcal{L} and denote it by $\text{com } \mathcal{L}$. The following results concerning relationships between commutativity and sharpness with informational completeness have been proven in⁴:

- Whenever $\dim \mathcal{H} \geq 2$ and $M \in \text{Obs}(\Sigma, \mathcal{H})$ is commutative, M is not informationally complete.
- A family of mutually commuting spectral measures is never informationally complete.
- If $P : \Sigma \rightarrow \mathcal{L}(\mathcal{H})$ is a spectral measure and $\rho \in \mathcal{S}(\mathcal{H})$, $[\rho]^P = \{\rho\}$ if and only if ρ is pure (a rank-1 projection) and there is an $X \in \Sigma$ such that $\rho = P(X)$.
- If \mathcal{O} is an informationally complete set of observables then $\dim \text{com } \mathcal{L} \leq 1$, where $\mathcal{L} = \bigcup_{M \in \mathcal{O}} \text{ran } M$.

The following are examples on informationally complete observables and sets of observables:

- The set $\{\mathbf{Q}_\theta\}_{\theta \in [0, \pi)}$ of the rotated quadratures introduced in Section I is informationally complete⁵ (Theorem 18.1)
- Equivalently with the above, the homodyne observable $\mathbf{G}_{\text{ht}} : \mathcal{B}([0, \pi) \times \mathbb{R}) \rightarrow \mathcal{L}(L^2(\mathbb{R}))$ defined by $\mathbf{G}_{\text{ht}}(\Theta \times X) = \pi^{-1} \int_{\Theta} \mathbf{Q}_\theta(X) d\theta$ is informationally complete.
- The covariant phase space observable \mathbf{G}_S introduced in Section I is informationally complete if and only if the support of the function $(q, p) \mapsto \text{tr}[SD(q, p)]$, i.e., the closure of the set of points $(q, p) \in \mathbb{R}^2$ such that $\text{tr}[SD(q, p)] \neq 0$, is \mathbb{R}^{25} .

A. Informational completeness within the set of pure states

Sometimes it is fruitful to consider informational completeness of an observable within a restricted set $\mathcal{P} \subseteq \mathcal{S}(\mathcal{H})$ of states; we are, e.g., already guaranteed that the pre-measurement state ρ is within \mathcal{P} and it is enough to only be able to discern between states in \mathcal{P} ⁶. Thus we

arrive at *informational completeness of an* $M \in \text{Obs}(\Sigma, \mathcal{H})$ *within* \mathcal{P} meaning that, whenever $\rho, \sigma \in \mathcal{P}$, $\rho \neq \sigma$, then $p_\rho^M \neq p_\sigma^M$. When the set \mathcal{P} consists of pure states, we identify it with $\{|\varphi\rangle\langle\varphi| \mid \varphi \in \mathcal{H}, \|\varphi\| = 1\}$; here, for any $\varphi \in \mathcal{H}$, we denote $[\varphi] := \{t\varphi \mid t \in \mathbb{T}\}$ and $\mathbb{T} := \{z \in \mathbb{C} \mid |z| = 1\}$. We get the following result for the case where we have to distinguish a pure state from other pure states:

Proposition 4. *Let $(\mathcal{H}_\oplus, \mathbf{P}_\oplus, J_\oplus)$ be the minimal Naimark dilation of Theorem 1 for an $M \in \text{Obs}(\Sigma, \mathcal{H})$ in a separable Hilbert space \mathcal{H} . The observable M is informationally complete within the set $\{|\varphi\rangle\langle\varphi| \mid \varphi \in \mathcal{H}, \|\varphi\| = 1\}$ of pure states if and only if $WJ_\oplus\varphi \notin J_\oplus\mathcal{H}$ whenever $\varphi \in \mathcal{H}$ and $W = \int_\Omega^\oplus W(x) d\mu(x) \in \mathcal{L}(\mathcal{H}_\oplus)$ is a decomposable isometry such that $WJ_\oplus\varphi \neq tJ_\oplus\varphi$ for all $t \in \mathbb{T}$.*

Proof. Let $\varphi, \psi \in \mathcal{H}$ be unit vectors. We have $p_{|\varphi\rangle\langle\varphi|}^M = p_{|\psi\rangle\langle\psi|}^M$ if and only if $\|(J_\oplus\varphi)(x)\| = \|(J_\oplus\psi)(x)\|$ for μ -a.a. $x \in \Omega$. One can construct measurable fields $x \mapsto \{e_n(x)\}_{n=1}^{m(x)} \subseteq \mathcal{H}(x)$, $x \mapsto \{f_n(x)\}_{n=1}^{m(x)} \subseteq \mathcal{H}(x)$ of orthonormal bases such that

$$e_1(x) = \begin{cases} \|(J_\oplus\varphi)(x)\|^{-1}(J_\oplus\varphi)(x), & (J_\oplus\varphi)(x) \neq 0 \\ \eta(x) & \text{otherwise} \end{cases},$$

$$f_1(x) = \begin{cases} \|(J_\oplus\psi)(x)\|^{-1}(J_\oplus\psi)(x), & (J_\oplus\psi)(x) \neq 0 \\ \eta(x) & \text{otherwise} \end{cases}$$

for μ -a.a. $x \in \Omega$, where $x \mapsto \eta(x) \in \mathcal{H}(x)$ is a measurable field of unit vectors. Defining $W(x) := \sum_{n=1}^{m(x)} |f_n(x)\rangle\langle e_n(x)|$ we may set up the decomposable isometry (even unitary) $W = \int_\Omega^\oplus W(x) d\mu(x)$ such that $J_\oplus\psi = WJ_\oplus\varphi$ if $p_{|\varphi\rangle\langle\varphi|}^M = p_{|\psi\rangle\langle\psi|}^M$ holds. In reverse, it is simple to check that, whenever W is a decomposable isometry such that $J_\oplus\psi = WJ_\oplus\varphi$, then $\|(J_\oplus\varphi)(x)\| = \|(J_\oplus\psi)(x)\|$ for μ -a.a. $x \in \Omega$, i.e., $p_{|\varphi\rangle\langle\varphi|}^M = p_{|\psi\rangle\langle\psi|}^M$. Thus, $p_{|\varphi\rangle\langle\varphi|}^M = p_{|\psi\rangle\langle\psi|}^M$ if and only if $J_\oplus\psi = WJ_\oplus\varphi$ with a decomposable isometry $W \in \mathcal{L}(\mathcal{H}_\oplus)$. The claim immediately follows from this observation and by noting that $|\varphi\rangle\langle\varphi| = |\psi\rangle\langle\psi|$ if and only if $J_\oplus\psi = tJ_\oplus\varphi$ for some $t \in \mathbb{T}$. \square

The above proposition implies the well-known fact stated earlier: a PVM in a separable Hilbert space cannot be informationally complete. In fact such a PVM \mathbf{P} is not informationally complete even within the set of pure states. Indeed, the isometry J_\oplus in the dilation of Theorem 1 for \mathbf{P} is unitary, i.e., $J_\oplus\mathcal{H} = \mathcal{H}_\oplus$.

For another example, as well known, the canonical phase Φ_{can} introduced in Section I is not informationally complete within the set of pure states. To see this using Proposition 4, let us give the minimal Naimark dilation of Theorem 1 for Φ_{can} in the form $(L^2([0, 2\pi), (2\pi)^{-1}d\theta), \mathbf{P}_{\text{can}}, J_{\text{can}})$, where \mathbf{P}_{can} is the canonical spectral measure of $L^2([0, 2\pi), (2\pi)^{-1}d\theta)$ and

$$J_{\text{can}} = \sum_{n=0}^{\infty} |\psi_n\rangle\langle n|, \quad \psi_n(\theta) = e^{-in\theta}, \quad 0 \leq \theta < 2\pi, \quad n \in \{0\} \cup \mathbb{N}.$$

Let $n \in \mathbb{N}$. Since $\psi_n(\theta) = e^{-in\theta}\psi_0(\theta)$ for all $\theta \in [0, 2\pi)$, defining the decomposable unitary operator W_n through $(W_n\psi)(\theta) = e^{-in\theta}\psi(\theta)$, $\psi \in L^2([0, 2\pi), (2\pi)^{-1}d\theta)$, $\theta \in [0, 2\pi)$, one has $J_{\text{can}}|n\rangle = \psi_n = W_n\psi_0 = W_nJ_{\text{can}}|0\rangle \neq tJ_{\text{can}}|0\rangle$. This proves the claim.

Let $\mathbf{M} \in \text{Obs}(\Sigma, \mathcal{H})$ be an observable in a separable Hilbert space \mathcal{H} with the minimal Naimark dilation $(\mathcal{H}_{\oplus}, \mathbf{P}_{\oplus}, J_{\oplus})$ of Theorem 1. Let us make a few observations and collect a couple conditions that guarantee that \mathbf{M} is not informationally complete within the set of pure states and a necessary and sufficient condition for this.

- *If there exist nonzero vectors $\varphi, \psi \in \mathcal{H}$ such that $\langle (J_{\oplus}\varphi)(x) | (J_{\oplus}\psi)(x) \rangle = 0$ for μ -almost all $x \in \Omega$ then \mathbf{M} is not informationally complete within the set of pure states.* To see this, fix φ and ψ satisfying the above condition. Without restricting generality, assume that $\|\varphi\| = 1 = \|\psi\|$. Hence, $\langle \varphi | \psi \rangle = \langle J_{\oplus}\varphi | J_{\oplus}\psi \rangle = 0$ and $\varphi_{\pm} = 2^{-1/2}(\varphi \pm \psi)$ are unit vectors for which $\varphi_+ \neq t\varphi_-$ for all $t \in \mathbb{T}$ and $\|(J_{\oplus}\varphi_+)(x)\| = \|(J_{\oplus}\varphi_-)(x)\|$ for μ -a.a. $x \in \Omega$, that is, $p_{|\varphi_+\rangle\langle\varphi_+|}^{\mathbf{M}} = p_{|\varphi_-\rangle\langle\varphi_-|}^{\mathbf{M}}$.

Note, however, that the existence of vectors φ and ψ of the above condition is not necessary for an informationally incomplete observable, a counterexample being the canonical phase: Assume that $\overline{(J_{\text{can}}\varphi)(\theta)}(J_{\text{can}}\psi)(\theta) = 0$ for $d\theta$ -a.a. $\theta \in [0, 2\pi)$. Then either $J_{\text{can}}\psi$ or $J_{\text{can}}\varphi$ is zero since any Hardy function vanishing on a set of positive measure is identically zero.

- *If there are disjoint sets $X_i \in \Sigma$ and nonzero vectors φ_i such that $\mathbf{M}(X_i)\varphi_i = \varphi_i$, $i = 1, 2$, then \mathbf{M} is not informationally complete within the set of pure states.* Indeed, let $\varphi_i \in \mathcal{H} \setminus \{0\}$ and X_i , $i = 1, 2$ be as above. Thus, $\mathbf{P}_{\oplus}(X_i)J_{\oplus}\varphi_i = J_{\oplus}\varphi_i$ for all $i = 1, 2$ implying that

$$|\langle (J_{\oplus}\varphi_1)(x) | (J_{\oplus}\varphi_2)(x) \rangle| \leq \|(J_{\oplus}\varphi_1)(x)\| \|(J_{\oplus}\varphi_2)(x)\| = 0$$

for μ -almost all $x \in \Omega$ so that \mathbf{M} is not informationally complete within the set of pure states.

- For any decomposable isometry $W = \int_{\Omega}^{\oplus} W(x) d\mu(x) \in \mathcal{L}(\mathcal{H}_{\oplus})$, define the operator $Z_W := J_{\oplus}^* W J_{\oplus}$. *The observable \mathbf{M} is informationally complete within the set of pure states if and only if, for any decomposable isometry $W \in \mathcal{L}(\mathcal{H}_{\oplus})$, the operator Z_W strictly decreases the norm (i.e., $\|Z_W \varphi\| < \|\varphi\|$) for any nonzero $\varphi \in \mathcal{H}$ such that $J_{\oplus} \varphi$ is not an eigenvector of W .* (Recall that an isometry may not have any eigenvalues and if eigenvalues exist they belong to \mathbb{T} .) To see this, note that, when $\varphi \in \mathcal{H}$ and $W \in \mathcal{L}(\mathcal{H}_{\oplus})$ is a decomposable isometry, the vector $W J_{\oplus} \varphi \neq t J_{\oplus} \varphi$ is not in the subspace $J_{\oplus} \mathcal{H} \cong \mathcal{H}$ if and only if its norm genuinely decreases under the ‘projection’ J_{\oplus}^* , i.e., $\|J_{\oplus}^* W J_{\oplus} \varphi\| < \|\varphi\|$. Thus we obtain the above as a reformulation of Proposition 4. Note that $\|Z_W\| \leq 1$ and, if $W J_{\oplus} \varphi = t J_{\oplus} \varphi$, then $Z_W \varphi = t \varphi$.

Suppose that $w : \Omega \rightarrow \mathbb{T}$ is a μ -measurable function and $W = \int_{\Omega}^{\oplus} w(x) I_{\mathcal{H}(x)} d\mu(x)$. Denote $Z_w := Z_W = \int_{\Omega} w(x) d\mathbf{M}(x)$. For the informational completeness of \mathbf{M} within the set of pure states, it is necessary that Z_w be strictly norm decreasing in the way defined above for any \mathbb{T} -valued measurable function w . We see immediately that this condition becomes also sufficient if \mathbf{M} is of rank 1. Note that, if w is not a constant on a set of positive measure, then the corresponding W does not have any eigenvalues. For example, in the case of the canonical phase, W_n ($n \neq 0$) does not have any eigenvalues and $Z_{W_n} = \int_0^{2\pi} e^{-in\theta} d\Phi_{\text{can}}(\theta) = \sum_{m=0}^{\infty} |m+n\rangle \langle m|$ is an isometry. Often we are interested in the state determination power of the rank-1 refinement of an observable which is why the rank-1 case is of particular importance.

Let us take a closer look at a couple of examples utilizing the observations made above.

Example 1. Consider a phase space observable G_S with some generating positive trace-1 operator S . Let us denote the closure of the range of S by \mathcal{K} . The dilation of Theorem 1 is given by the Hilbert space $L^2(\mathbb{R}^2) \otimes \mathcal{K}$ identified here with the corresponding L^2 -space of (equivalence classes of) \mathcal{K} -valued functions, the canonical spectral measure $\mathbf{P}_{\oplus} : \mathcal{B}(\mathbb{R}^2) \rightarrow \mathcal{L}(L^2(\mathbb{R}^2) \otimes \mathcal{K})$, $(\mathbf{P}_{\oplus}(Z)\eta)(q, p) = \chi_Z(q, p)\eta(q, p)$ for all $Z \in \mathcal{B}(\mathbb{R}^2)$, $\eta \in L^2(\mathbb{R}^2) \otimes \mathcal{K}$, and a.a. $(q, p) \in \mathbb{R}^2$, and the isometry $J_{\oplus} : L^2(\mathbb{R}) \rightarrow L^2(\mathbb{R}^2) \otimes \mathcal{K}$,

$$(J_{\oplus} \varphi)(q, p) = \frac{1}{\sqrt{2\pi}} S^{1/2} D(q, p)^* \varphi, \quad \varphi \in L^2(\mathbb{R}), \quad (q, p) \in \mathbb{R}^2.$$

It follows that \mathbf{G}_S is informationally complete within the set of pure states if and only if, whenever $\mathbb{R}^2 \ni (q, p) \mapsto W(q, p) \in \mathcal{L}(\mathcal{K})$ is a weakly measurable field of isometries, the operator

$$Z_W = \frac{1}{2\pi} \int_{\mathbb{R}^2} D(q, p) S^{1/2} W(q, p) S^{1/2} D(q, p)^* dq dp$$

strictly decreases the norm of any nonzero vector φ such that $J_{\oplus} \varphi$ is not an eigenvector of W . Especially, if $w : \mathbb{R}^2 \rightarrow \mathbb{T}$ is measurable then $Z_w = \int_{\mathbb{R}^2} w(q, p) d\mathbf{G}_S(q, p)$ is strictly norm decreasing in the above sense if \mathbf{G}_S is informationally complete (within the set of pure states). If S is of rank 1 (i.e., \mathbf{G}_S is rank-1) then \mathbf{G}_S is informationally complete within the set of pure states if and only if the operators Z_w are strictly norm decreasing as above.

Let $S = \sum_{i=1}^{\text{rank } S} s_i |\varphi_i\rangle\langle\varphi_i|$ be the spectral decomposition of S where $\varphi_i \in L^2(\mathbb{R})$ is a unit eigenvector associated to the eigenvalue $s_i \in (0, 1]$ (and $\langle\varphi_i|\varphi_j\rangle = \delta_{ij}$, $\text{tr}[S] = \sum_i s_i = 1$). Pick a representative $\mathbb{R} \ni x \mapsto \varphi_i(x) \in \mathbb{C}$ from each class φ_i such that $\sum_{i=1}^{\text{rank } S} s_i |\varphi_i(x)|^2 < \infty$ for all $x \in \mathbb{R}$ and define a positive semidefinite integral kernel $K_S : \mathbb{R}^2 \rightarrow \mathbb{C}$ by

$$K_S(x, y) := \sum_{i=1}^{\text{rank } S} s_i \varphi_i(x) \overline{\varphi_i(y)}, \quad (x, y) \in \mathbb{R}^2.$$

Indeed, $|K_S(x, y)|^2 \leq K_S(x, x)K_S(y, y)$ and $\int_{\mathbb{R}} K_S(x, x) dx = 1$ by the monotone convergence theorem. Now $X_{K_S} := \{x \in \mathbb{R} \mid K_S(x, x) \neq 0\}$ is essentially unique in the sense that, if \tilde{K}_S is another integral kernel of S then X_{K_S} and $X_{\tilde{K}_S}$ differ in the set of Lebesgue measure zero. For any $\varphi, \psi \in L^2(\mathbb{R})$ and $(q, p) \in \mathbb{R}^2$ one gets

$$\begin{aligned} 2\pi \langle (J_{\oplus} \varphi)(q, p) | (J_{\oplus} \psi)(q, p) \rangle &= \langle D(q, p)^* \varphi | S D(q, p)^* \psi \rangle \\ &= \int_{X_{K_S}} \int_{X_{K_S}} \overline{(D(q, p)^* \varphi)(x)} S(x, y) (D(q, p)^* \psi)(y) dx dy \\ &= \int \int_{Y_{S, \varphi, \psi}^q} e^{ipx} \overline{\varphi(x+q)} K_S(x, y) e^{-ipy} \psi(y+q) dx dy \end{aligned}$$

where $Y_{S, \varphi, \psi}^q = X_{K_S} \times X_{K_S} \cap \{x \mid \varphi(x+q) \neq 0\} \times \{y \mid \psi(y+q) \neq 0\}$. If there exists an $R > 0$ such that $X_{K_S} \setminus [-R, R]$ is of measure zero (e.g. $S = |\chi_{[0,1]}\rangle\langle\chi_{[0,1]}|$) then it is easy to find φ and ψ such that $Y_{S, \varphi, \psi}^q$ is zero measurable for all $q \in \mathbb{R}$ and, hence, \mathbf{G}_S is not informationally complete within the set of pure states by above observation.

To connect our analysis with earlier results, define a continuous square-integrable function

$$\hat{S} : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad (q, p) \mapsto \hat{S}(q, p) := \text{tr}[D(q, p)S] = e^{iqp/2} \int_{\mathbb{R}} e^{ipx} K_S(x, x+q) dx.$$

If \hat{S} is integrable, $K_S(x, x + q) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{-iqp/2} e^{-ipx} \hat{S}(q, p) dp$ for all $q \in \mathbb{R}$ and a.a. $x \in \mathbb{R}$. If, additionally, $X_{K_S} \setminus [-R, R]$ is of measure zero for some $R > 0$, $\hat{S}(q, p) = 0$ for all $p \in \mathbb{R}$ if $|q| > 2R$ but \hat{S} need not be compactly supported (e.g. $S = |\chi_{[0,1]}\rangle\langle\chi_{[0,1]}|$ for which $\hat{S}(0, p) = i(1 - e^{ip})/p$ for all $p \neq 0$). Assume then that the support of \hat{S} is compact and thus contained in a rectangle $[-R_0, R_0] \times [-R_0, R_0]$. Now \hat{S} is integrable and $S(x, x + q) = 0$ for (almost) all x and q such that $|q| > R_0$. Immediately one finds unit vectors $\varphi, \psi \in L^2(\mathbb{R})$ such that $\langle (J_{\oplus} \varphi)(q, p) | (J_{\oplus} \psi)(q, p) \rangle = 0$ for all $(q, p) \in \mathbb{R}^2$ thus showing that \mathbf{G}_S is not informationally complete within the set of pure states. Hence, we have obtained Proposition 20(c) of⁶ as a special case.

Example 2. Let $\mathbf{M} : 2^\Omega \rightarrow \mathcal{L}(\mathcal{H})$ be an observable with an at most countably infinite value space $\Omega = \{x_1, x_2, \dots\}$, and denote $\mathbf{M}_i := \mathbf{M}(\{x_i\})$ for all $i = 1, 2, \dots$. Denote the closure of the range of \mathbf{M}_i by \mathcal{K}_i for each i and define the Hilbert space $\mathcal{K} := \bigoplus_i \mathcal{K}_i$ which is equipped with the canonical spectral measure $\mathbf{P} : 2^\Omega \rightarrow \mathcal{L}(\mathcal{K})$ defined by $\mathbf{P}(\{x_i\}) \bigoplus_j \varphi_j := \varphi_i$ for all i and all $\bigoplus_j \varphi_j \in \mathcal{K}$. Moreover, define the isometry $J : \mathcal{H} \rightarrow \mathcal{K}$, $\varphi \mapsto J\varphi = \bigoplus_i \mathbf{M}_i^{1/2} \varphi$. The triple $(\mathcal{K}, \mathbf{P}, J)$ is a minimal Naimark dilation for \mathbf{M} like the one presented in Theorem 1.

The observable \mathbf{M} is informationally complete within the set of pure states if and only if, for any isometries $W_i \in \mathcal{L}(\mathcal{K}_i)$ and any $\varphi \in \mathcal{H}$ such that there is no $t \in \mathbb{T}$ such that $W_i \mathbf{M}_i^{1/2} \varphi = t \mathbf{M}_i^{1/2} \varphi$ for all i , one has $\|Z_W \varphi\| < \|\varphi\|$ where $Z_W := \sum_i \mathbf{M}_i^{1/2} W_i \mathbf{M}_i^{1/2}$. This is a direct consequence of our earlier observations by noting that, when $W_i \in \mathcal{L}(\mathcal{K}_i)$ are isometries and $W := \bigoplus_i W_i$, then $WJ\varphi = tJ\varphi$ for some $\varphi \in \mathcal{H}$ and $t \in \mathbb{T}$ if and only if $W_i \mathbf{M}_i^{1/2} \varphi = t \mathbf{M}_i^{1/2} \varphi$ for all i . If \mathbf{M} is of rank 1, this condition can be simplified: \mathbf{M} is informationally complete within the set of pure states if and only if, for any (nonconstant) function $i \mapsto w_i \in \mathbb{T}$ and any nonzero $\varphi \in \mathcal{H}$ (such that $w_i \mathbf{M}_i^{1/2} \varphi \neq t \mathbf{M}_i^{1/2} \varphi$ for all i), one has $\|Z_w \varphi\| < \|\varphi\|$, where $Z_w = \sum_i w_i \mathbf{M}_i$.

Finally, we note that, if $\{\varphi_{ik}\}_{k=1}^{\dim \mathcal{K}_i}$ is an orthonormal basis of \mathcal{K}_i for each i one can define (linearly independent) vectors $d_{ik} := \mathbf{M}_i^{1/2} \varphi_{ik}$, $k < \dim \mathcal{K}_i + 1$, such that $\mathbf{M}_i^{1/2} = \sum_k |\varphi_{ik}\rangle\langle\varphi_{ik}| \mathbf{M}_i^{1/2} = \sum_k |\varphi_{ik}\rangle\langle d_{ik}|$, $J = \sum_i \sum_k |e_{ik}\rangle\langle d_{ik}|$, and $\mathbf{P}_i := \mathbf{P}(\{x_i\}) = \sum_k |e_{ik}\rangle\langle e_{ik}|$, where $e_{ik} := \bigoplus_j \delta_{ji} \varphi_{ik}$; compare to Section I.

VII. EXTREME OBSERVABLES

The relevant mathematical structures in quantum theory, sets of states, observables, channels, and instruments, are convex. For example, for observables $\mathbf{M}, \mathbf{M}' \in \text{Obs}(\Sigma, \mathcal{H})$ and $p \in [0, 1]$, one can determine a mixed observable, a convex combination $p\mathbf{M} + (1 - p)\mathbf{M}'$ by

$$(p\mathbf{M} + (1 - p)\mathbf{M}')(X) = p\mathbf{M}(X) + (1 - p)\mathbf{M}'(X), \quad X \in \Sigma.$$

Such mixing of devices can be seen as classical noise produced by an imprecise implementation that produces a measurement of \mathbf{M} with relative frequency p and something else otherwise.

An element $x \in K$ in a convex set K is called *extreme* if, for any $y, z \in K$ and $p \in (0, 1)$, $x = py + (1 - p)z$ implies $x = y = z$. Thus extreme quantum devices are free of classical noise due to mixing. The extreme elements of the set of states $\mathcal{S}(\mathcal{H})$ are the rank-1 projections $|\varphi\rangle\langle\varphi|$, $\varphi \in \mathcal{H}$, $\|\varphi\| = 1$, called as *pure states*, whereas the extreme effects are projections. The general characterizations of extremality for quantum devices follow ultimately from the following result¹. Below, we denote by $\mathbf{CP}(\mathcal{A}; \mathcal{H})$ the set of unital completely positive maps $\Phi : \mathcal{A} \rightarrow \mathcal{L}(\mathcal{H})$ where \mathcal{A} is a unital C^* -algebra and \mathcal{H} is a Hilbert space.

Theorem 8. *Let $\Phi \in \mathbf{CP}(\mathcal{A}; \mathcal{H})$ and pick a minimal Stinespring dilation (\mathcal{M}, π, J) for Φ . The map Φ is an extreme point of the convex set $\mathbf{CP}(\mathcal{A}; \mathcal{H})$ if and only if the map*

$$(\text{ran } \pi)' \ni D \mapsto J^*DJ \in \mathcal{L}(\mathcal{H})$$

defined on the commutant of the range of π is an injection.

Suppose now that \mathcal{A} is a von Neumann algebra and denote the set of normal maps within $\mathbf{CP}(\mathcal{A}; \mathcal{H})$ by $\mathbf{NCP}(\mathcal{A}; \mathcal{H})$. The convex subset $\mathbf{NCP}(\mathcal{A}; \mathcal{H})$ is a facet of $\mathbf{CP}(\mathcal{A}; \mathcal{H})$ meaning that, if $\Phi = t\Phi_1 + (1 - t)\Phi_2 \in \mathbf{NCP}(\mathcal{A}; \mathcal{H})$ for some $\Phi_1, \Phi_2 \in \mathbf{CP}(\mathcal{A}; \mathcal{H})$, then $\Phi_1, \Phi_2 \in \mathbf{NCP}(\mathcal{A}; \mathcal{H})$ as well. From this it follows that the extreme elements of $\mathbf{NCP}(\mathcal{A}; \mathcal{H})$ are characterized in exactly the same way as in Theorem 8.

We usually say shortly that an observable $\mathbf{M} : \Sigma \rightarrow \mathcal{L}(\mathcal{H})$ is extreme if \mathbf{M} is an extreme element of $\text{Obs}(\Sigma, \mathcal{H})$. We may elaborate the above extremality characterization in the case of quantum observables³¹.

Theorem 9. Let $(\mathcal{H}_\oplus, \mathbf{P}_\oplus, J_\oplus)$ be the minimal Naimark dilation of Theorem 1 for an $\mathbf{M} \in \text{Obs}(\Sigma, \mathcal{H})$ in a separable Hilbert space \mathcal{H} . The observable \mathbf{M} is extreme if and only if, for any decomposable operator $D = \int_\Omega^\oplus D(x) d\mu(x) \in \mathcal{L}(\mathcal{H}_\oplus)$, the condition $J_\oplus^* D J_\oplus = 0$ implies $D = 0$.

It is an immediate result of Theorem 9 that PVMs are extreme. This can also be proven directly by using the fact that projections are the extreme elements of the set of effects. Also, if (Ω, Σ) is nice and $\dim \mathcal{H} < \infty$ then an extreme observable $\mathbf{M} \in \text{Obs}(\Sigma, \mathcal{H})$ is discrete. Indeed, using an exactly measurable function $f : \Omega \rightarrow \mathbb{R}$ such that $f(\Omega) \in \mathcal{B}(\mathbb{R})$, we now obtain an extreme observable $\mathbf{M} \circ f^{-1} \in \text{Obs}(\mathcal{B}(\mathbb{R}), \mathcal{H})$ which is supported on an at most countable set $\{\lambda_1, \lambda_2, \dots\} \subset \mathbb{R}^{15}$ (Section 5); see also⁷. This means that \mathbf{M} is supported by the set $\bigcup_i f^{-1}(\{\lambda_i\})$ where $f^{-1}(\{\lambda_i\})$ are atoms of Σ . Below are some examples on extreme observables which are not PVMs.

- One can show that the phase space observable \mathbf{G}_S introduced in Section I is extreme if and only if S is pure, $S = |\psi\rangle\langle\psi|$, and $(q, p) \mapsto \langle\psi|D(q, p)\psi\rangle \neq 0$ for all $(q, p) \in \mathbb{R}^{220}$. Hence, if \mathbf{G}_S is extreme then it is informationally complete (but the converse does not hold). Especially, when $S = |0\rangle\langle 0|$, i.e., the generating state is the vacuum state, then we get the rank-1 informationally complete extreme observable $\mathbf{G}_{|0\rangle\langle 0|}$.
- The canonical phase Φ_{can} introduced in Section I is extreme¹⁹.
- Fix a number $m > 0$ and denote by $\hat{\varphi}$ the Fourier-Plancherel transformation of $\varphi \in L^2(\mathbb{R})$; if $\varphi \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ we may write

$$\hat{\varphi}(p) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-ipx} \varphi(x) dx, \quad p \in \mathbb{R}.$$

The canonical time-of-arrival observable $\tau : \mathcal{B}(\mathbb{R}) \rightarrow \mathcal{L}(L^2(\mathbb{R}))$ for a free mass- m particle moving in \mathbb{R} defined through

$$\langle\varphi|\tau(X)\psi\rangle = \frac{1}{2\pi m} \int_X \int_0^\infty \int_0^\infty e^{\frac{it}{m}(p_2^2 - p_1^2)} (\overline{\hat{\varphi}(p_1)} \hat{\psi}(p_2) + \overline{\hat{\varphi}(-p_1)} \hat{\psi}(-p_2)) \sqrt{p_1 p_2} dp_1 dp_2 dt$$

for any $X \in \mathcal{B}(\mathbb{R})$ and any vectors φ and ψ from the Schwartz space of rapidly decreasing functions, is extreme¹⁷.

We see from Theorem 9 that *if \mathbf{M} is extreme so is its rank-1 refinement \mathbf{M}^{132}* . The two first examples given above are of rank 1. The third example, however, is of rank 2. The

rank-1 refinement of the canonical time observable is $\tau^1 : 2^{\{1,2\}} \otimes \mathcal{B}(\mathbb{R}) \rightarrow \mathcal{L}(L^2(\mathbb{R}))$,

$$\langle \varphi | \tau^1(\{k\} \times X) \psi \rangle = \frac{1}{2\pi m} \int_X \int_0^\infty \int_0^\infty e^{\frac{it}{2m}(p_2^2 - p_1^2)} \overline{\hat{\varphi}((-1)^k p_1)} \hat{\psi}((-1)^k p_2) \sqrt{p_1 p_2} dp_1 dp_2 dt$$

for all $X \in \mathcal{B}(\mathbb{R})$, $k = 1, 2$, and all φ and ψ from the Schwartz space of rapidly decreasing functions. Thus, τ^1 is extreme too.

Let us next consider an example where we show how to construct an observable in a separable Hilbert space with all the optimality properties discussed this far, i.e., a post-processing clean (rank-1) informationally complete extreme observable.

Example 3. Let $\mathbb{N}_\infty := \mathbb{N} \cup \{\infty\}$ and $\mathbb{N}_d := \{1, 2, \dots, d\}$ for each $d \in \mathbb{N}$. Let \mathcal{H}_d , $d \in \mathbb{N}_\infty$, be a d -dimensional Hilbert space with \mathcal{H}_∞ having the orthonormal basis $\{|n\rangle\}_{n=1}^\infty$. We assume that $\mathcal{H}_{d_1} \subseteq \mathcal{H}_{d_2} \subseteq \mathcal{H}_\infty$ whenever $d_1 \leq d_2$ so that, for any $d \in \mathbb{N}$, $\{|n\rangle\}_{n=1}^d$ is an orthonormal basis for \mathcal{H}_d .

For all $n, m \in \mathbb{N}$, pick some numbers $p_{nm} > 0$ for which

$$p := \sum_{n,m=1}^\infty p_{nm} < \infty.$$

Let $n, m \in \mathbb{N}$ such that $n < m$. Define the following vectors:

$$f_{nn} := |n\rangle, \quad f_{nm} := |n\rangle + |m\rangle, \quad f_{mn} := |n\rangle - i|m\rangle,$$

so that

$$\begin{aligned} |f_{nm}\rangle \langle f_{nm}| &= |n\rangle \langle n| + |n\rangle \langle m| + |m\rangle \langle n| + |m\rangle \langle m|, \\ |f_{mn}\rangle \langle f_{mn}| &= |n\rangle \langle n| + i|n\rangle \langle m| - i|m\rangle \langle n| + |m\rangle \langle m| \end{aligned}$$

and thus $|n\rangle \langle n| = |f_{nn}\rangle \langle f_{nn}|$,

$$\begin{aligned} 2|n\rangle \langle m| &= (|f_{nm}\rangle \langle f_{nm}| - |f_{nn}\rangle \langle f_{nn}| - |f_{mm}\rangle \langle f_{mm}|) \\ &\quad - i(|f_{mn}\rangle \langle f_{mn}| - |f_{nn}\rangle \langle f_{nn}| - |f_{mm}\rangle \langle f_{mm}|), \\ 2|m\rangle \langle n| &= (|f_{nm}\rangle \langle f_{nm}| - |f_{nn}\rangle \langle f_{nn}| - |f_{mm}\rangle \langle f_{mm}|) \\ &\quad + i(|f_{mn}\rangle \langle f_{mn}| - |f_{nn}\rangle \langle f_{nn}| - |f_{mm}\rangle \langle f_{mm}|). \end{aligned}$$

Hence, for any $d \in \mathbb{N}_\infty$, the linearly independent set

$$\mathcal{B}_d := \{p_{nm} |f_{nm}\rangle \langle f_{nm}| \in \mathcal{L}(\mathcal{H}_d) \mid n, m < d + 1\}$$

has d^2 elements and is a basis of $\mathcal{L}(\mathcal{H}_d)$. If $d = \infty$ then the linear span of \mathcal{B}_∞ is dense in $\mathcal{L}(\mathcal{H}_\infty)$ with respect to the weak operator topology.

For each $\mathcal{I} \subseteq \mathbb{N}_d \times \mathbb{N}_d$ (or $\mathcal{I} \subseteq \mathbb{N} \times \mathbb{N}$ if $d = \infty$) we define a positive trace-class operator

$$S_{\mathcal{I}} := \sum_{(n,m) \in \mathcal{I}} p_{nm} |f_{nm}\rangle \langle f_{nm}| \in \mathcal{L}(\mathcal{H}_d).$$

Indeed, $\|S_{\mathcal{I}}\| \leq \text{tr} \|S_{\mathcal{I}}\| = \text{tr} [S_{\mathcal{I}}] = \sum_{(n,m) \in \mathcal{I}} p_{nm} \|f_{nm}\|^2 < 2p$.

Let $\mathcal{I} \subseteq \mathbb{N}_d \times \mathbb{N}_d$ (or $\mathcal{I} \subseteq \mathbb{N} \times \mathbb{N}$ if $d = \infty$), and let $\mathcal{I}_0 \subseteq \mathcal{I}$ be such that the vectors $f_{nm} \in \mathcal{H}_d$, $(n, m) \in \mathcal{I}_0$, form a basis of the vector space $\text{lin}\{f_{nm} \mid (n, m) \in \mathcal{I}\} \subseteq \mathcal{H}_d$ whose closure is the range of $S_{\mathcal{I}}$. Now the maximal number of linearly independent elements of $\{f_{nm} \mid (n, m) \in \mathcal{I}\} \subseteq \mathcal{H}_d$ is $\#\mathcal{I}_0 \leq d$. Hence, the rank of $S_{\mathcal{I}}$ (and $S_{\mathcal{I}_0} \leq S_{\mathcal{I}}$) is $\#\mathcal{I}_0$ and

$$S_{\mathcal{I}} = \sum_{k=1}^{\#\mathcal{I}_0} |\varphi_k^{\mathcal{I}}\rangle \langle \varphi_k^{\mathcal{I}}|,$$

where the eigenvectors $\varphi_k^{\mathcal{I}} \in \mathcal{H}_d$ form an orthogonal set and the eigenvalues $\|\varphi_k^{\mathcal{I}}\|^2 > 0$ are such that $\sum_{k=1}^{\#\mathcal{I}_0} \|\varphi_k^{\mathcal{I}}\|^2 = \text{tr} [S_{\mathcal{I}}] < \infty$. Note that

$$\varphi_k^{\mathcal{I}} = \|\varphi_k^{\mathcal{I}}\|^{-2} S_{\mathcal{I}} \varphi_k^{\mathcal{I}} = \|\varphi_k^{\mathcal{I}}\|^{-2} \sum_{(n,m) \in \mathcal{I}} p_{nm} \langle f_{nm} | \varphi_k^{\mathcal{I}} \rangle f_{nm} \in \text{ran } S_{\mathcal{I}} = \text{ran } S_{\mathcal{I}_0} = \overline{\text{lin}\{f_{nm} \mid (n, m) \in \mathcal{I}_0\}}$$

and $\sum_{k=1}^{\#\mathcal{I}_0} \|\varphi_k^{\mathcal{I}}\|^{-2} |\varphi_k^{\mathcal{I}}\rangle \langle \varphi_k^{\mathcal{I}}|$ is the projection from \mathcal{H}_d onto the $\#\mathcal{I}_0$ -dimensional Hilbert space $\text{ran } S_{\mathcal{I}_0}$.

We assume next that $\#\mathcal{I}_0 = d$ so that $\{f_{nm} \mid (n, m) \in \mathcal{I}_0\}$ is a basis of \mathcal{H}_d and $S_{\mathcal{I}}$ is of full rank and thus invertible. An example of such a situation is, $\mathcal{I} = \mathbb{N}_d \times \mathbb{N}_d$ and $\mathcal{I}_0 = \{(n, n) \mid n \in \mathbb{N}_d\}$ (if $d < \infty$). Let $2^{\mathcal{I}}$ be the power set of \mathcal{I} and define the following discrete rank-1 POVM $\mathbf{M} : 2^{\mathcal{I}} \rightarrow \mathcal{L}(\mathcal{H}_d)$:

$$\mathbf{M}_{nm} := |\sqrt{p_{nm}} S_{\mathcal{I}}^{-1/2} f_{nm}\rangle \langle \sqrt{p_{nm}} S_{\mathcal{I}}^{-1/2} f_{nm}|, \quad (n, m) \in \mathcal{I}.$$

Since, for any complex numbers c_{nm} such that $\sup\{|c_{nm}| \mid (n, m) \in \mathcal{I}\} < \infty$,

$$\sum_{(n,m) \in \mathcal{I}} c_{nm} \mathbf{M}_{nm} = \sum_{(n,m) \in \mathcal{I}} c_{nm} p_{nm} |S_{\mathcal{I}}^{-1/2} f_{nm}\rangle \langle S_{\mathcal{I}}^{-1/2} f_{nm}| = 0$$

if and only if $\sum_{(n,m) \in \mathcal{I}} c_{nm} p_{nm} |f_{nm}\rangle \langle f_{nm}| = S_{\mathcal{I}}^{1/2} \sum_{(n,m) \in \mathcal{I}} c_{nm} p_{nm} |S_{\mathcal{I}}^{-1/2} f_{nm}\rangle \langle S_{\mathcal{I}}^{-1/2} f_{nm}| S_{\mathcal{I}}^{1/2} = 0$ if and only if $c_{nm} \equiv 0$, the observable \mathbf{M} is extreme with $N = \#\mathcal{I} \geq \#\mathcal{I}_0 = d$ elements. Automatically, $d \leq N \leq d^2$ which must hold for any extreme rank-1 POVM. If $\mathcal{I} = \mathbb{N}_d \times \mathbb{N}_d$

(or $\mathcal{I} = \mathbb{N} \times \mathbb{N}$ if $d = \infty$) then $N = d^2$ and \mathcal{B}_d is a basis of $\mathcal{L}(\mathcal{H}_d)$ showing that \mathbf{M} is also informationally complete. Note that, in the case $N = d$, we get a PVM \mathbf{M} . When N runs from d to d^2 the value determination ability weakens but state determination power increases.

A. Joint measurements of extreme observables

We now concentrate on joint measurements involving extreme observables. Theorem 2 together with the extremality characterizations of theorems 8 and 9 imply that, if \mathbf{M} is an extreme observable and \mathbf{M}' is an observable jointly measurable with \mathbf{M} and Φ is an \mathbf{M} -channel, there is a unique joint observable for \mathbf{M} and \mathbf{M}' and a unique joint instrument for \mathbf{M} and Φ . As proven in¹⁶, we obtain the following results:

Theorem 10. *Let observables $\mathbf{M} : \Sigma \rightarrow \mathcal{L}(\mathcal{H})$ and $\mathbf{M}' : \Sigma' \rightarrow \mathcal{L}(\mathcal{H})$ be jointly measurable, and let $\Phi : \mathcal{L}(\mathcal{K}) \rightarrow \mathcal{L}(\mathcal{H})$ be a channel compatible with \mathbf{M} . If \mathbf{M} is extreme then*

(i) *the \mathbf{M} -instrument $\mathcal{J} : \mathcal{L}(\mathcal{K}) \times \Sigma \rightarrow \mathcal{L}(\mathcal{H})$ such that $\mathcal{J}(B, \Omega) = \Phi(B)$ for all $B \in \mathcal{L}(\mathcal{K})$ is unique and*

(ii) *the joint observable $\mathbf{N} : \Sigma \otimes \Sigma' \rightarrow \mathcal{L}(\mathcal{H})$ for \mathbf{M} and \mathbf{M}' is unique.*

Moreover, if \mathbf{M} is a PVM,

(i)' *$\mathbf{M}(X)\Phi(B) = \Phi(B)\mathbf{M}(X)$ for all $X \in \Sigma$ and all $B \in \mathcal{L}(\mathcal{K})$ and the only \mathbf{M} -instrument $\mathcal{J} : \mathcal{L}(\mathcal{K}) \times \Sigma \rightarrow \mathcal{L}(\mathcal{H})$ such that $\mathcal{J}(\cdot, \Omega) = \Phi$ is given by*

$$\mathcal{J}(B, X) = \Phi(B)\mathbf{M}(X) = \mathbf{M}(X)^{1/2}\Phi(B)\mathbf{M}(X)^{1/2}, \quad B \in \mathcal{L}(\mathcal{K}), \quad X \in \Sigma,$$

and

(ii)' *$\mathbf{M}(X)\mathbf{M}'(Y) = \mathbf{M}'(Y)\mathbf{M}(X)$ for all $X \in \Sigma$ and all $Y \in \Sigma'$ and the only joint observable $\mathbf{N} : \Sigma \otimes \Sigma' \rightarrow \mathcal{L}(\mathcal{H})$ for \mathbf{M} and \mathbf{M}' is determined by*

$$\mathbf{N}(X \times Y) = \mathbf{M}(X)\mathbf{M}'(Y) = \mathbf{M}(X)^{1/2}\mathbf{M}'(Y)\mathbf{M}(X)^{1/2}, \quad X \in \Sigma, \quad Y \in \Sigma'.$$

The above result essentially means that there is only one way in which an extreme observable can be measured if we fix the unconditioned state transformation associated with

the measurement. Similarly, there is only one observable incorporating an extreme marginal observable and some other fixed observable. The corresponding conditions for joint measurements involving PVMs, being from a special subclass of extreme observables, are even more stringent: compatibility or joint measurability with a PVM requires that the other measurement device (observable or channel) commutes with the PVM.

VIII. OBSERVABLES DETERMINING THEIR VALUES

We say that an observable $M : \Sigma \rightarrow \mathcal{L}(\mathcal{H})$ determines its values if for any set of its outcomes we may prepare the system into a state such that, in a measurement of M , the values obtained are approximately localized within the given set. Formally, this means that, for any $X \in \Sigma$ such that $M(X) \neq 0$ and any $\varepsilon \in (0, 1]$, there is a state $\rho \in \mathcal{S}(\mathcal{H})$ such that $p_\rho^M(X) > 1 - \varepsilon$.

Suppose that $M : \Sigma \rightarrow \mathcal{L}(\mathcal{H})$ determines its values and $X \in \Sigma$ is such that $M(X) > 0$. We may evaluate for any $\varepsilon \in (0, 1]$

$$\|M(X)\| = \sup_{\rho \in \mathcal{S}(\mathcal{H})} \text{tr}[\rho M(X)] = \sup_{\rho \in \mathcal{S}(\mathcal{H})} p_\rho^M(X) > 1 - \varepsilon$$

showing that $\|M(X)\| = 1$. Indeed, we see that this reasoning can easily be inverted: an observable determines its values if and only if it has the *norm-1 property*, i.e., $\|M(X)\| = 1$ for all $X \in \Sigma$ such that $M(X) \neq 0$.

A more stringent condition than the norm-1 property is the *eigenvalue-1 property*: $M \in \text{Obs}(\Sigma, \mathcal{H})$ is an *eigenvalue-1* observable if and only if, whenever $X \in \Sigma$ is such that $M(X) \neq 0$, then $M(X)$ has the eigenvalue 1. This means that for any $X \in \Sigma$ such that $M(X) \neq 0$ there is a state $\rho_X \in \mathcal{S}(\mathcal{H})$ “localized” in X in the sense that $p_{\rho_X}^M(X) = 1$, i.e., the approximate “ ε -localization” associated with norm-1 observables can be replaced with exact localization. Of course, we may always assume that ρ_X above is pure. Clearly, PVMs are eigenvalue-1 observables and there exist norm-1 POVMs which have not eigenvalue-1 property, e.g. the canonical phase Φ_{can} .

Consider then a norm-1 observable $M \in \text{Obs}(\Sigma, \mathcal{H})$ in a finite-dimensional Hilbert space. Denote $d = \dim \mathcal{H}$. Since any effect has fully discrete spectrum in finite dimensions, M has the eigenvalue-1 property. For $i = 1, 2$, let the sets $X_i \in \Sigma$ and unit vectors $\varphi_i \in \mathcal{H}$ be such that $M(X_i)\varphi_i = \varphi_i$ and $X_1 \cap X_2 = \emptyset$ (we assume that M is not trivial). By using the minimal

Naïmark dilation for \mathbf{M} , one gets $\langle \varphi_1 | \varphi_2 \rangle = 0$. Hence, there exist at most d pairwise disjoint sets $X \in \Sigma$ such that $\mathbf{M}(X) \neq 0$. If (Ω, Σ) is, e.g., nice we may conclude that \mathbf{M} is discrete, i.e. there exist points $x_i \in \Omega$, $i = 1, \dots, N \leq d$, so that $\mathbf{M} = \sum_{i=1}^N \mathbf{M}(A_{x_i}) \delta_{x_i}$ where A_{x_i} is the atom associated with x_i . This result follows by using an exactly measurable function $f : \Omega \rightarrow \mathbb{R}$ such that $f(\Omega) \in \mathcal{B}(\mathbb{R})$ and results of¹⁵ (Section 5).

Finally, let us recall that an eigenvalue-1 observable cannot be informationally complete even within the set of pure states. Indeed, it is easy to show that this holds in arbitrary (nonseparable) Hilbert spaces. The question, whether a norm-1 observable can be informationally complete, remains to be answered.

IX. PRE-PROCESSING AND PRE-PROCESSING MAXIMALITY

Let us start this section with a definition.

Definition 7. Let (Ω, Σ) be a measurable space and \mathcal{H} and \mathcal{H}' be Hilbert spaces. We say that an observable $\mathbf{M}' : \Sigma \rightarrow \mathcal{L}(\mathcal{H}')$ is a *pre-processing* of an observable $\mathbf{M} : \Sigma \rightarrow \mathcal{L}(\mathcal{H})$ if there is a channel $\Phi : \mathcal{L}(\mathcal{H}) \rightarrow \mathcal{L}(\mathcal{H}')$ such that $\mathbf{M}'(X) = \Phi(\mathbf{M}(X))$ for all $X \in \Sigma$.

The above definition means that $p_\sigma^{\mathbf{M}'} = p_{\Phi_*(\sigma)}^{\mathbf{M}}$ for all $\sigma \in \mathcal{S}(\mathcal{H}')$, that is, we may measure \mathbf{M}' by first transforming the system with the channel Φ and then measuring \mathbf{M} . The predual channel Φ_* is here seen as a form of quantum pre-processing that is used to process the incoming state carrying quantum information before the measurement of \mathbf{M} .

Pre-processing gives naturally rise to a partial order within the class of observables with the fixed value space (Ω, Σ) and varying system's Hilbert space \mathcal{H}^3 . We denote $\mathbf{M}' \leq_{\text{pre}} \mathbf{M}$ if \mathbf{M}' is a pre-processing of \mathbf{M} by some channel. We may thus ask which are the pre-processing maximal or clean observables. Maximality of an $\mathbf{M} \in \text{Obs}(\Sigma, \mathcal{H})$ means that if $\mathbf{M} \leq_{\text{pre}} \mathbf{M}'$ with some observable $\mathbf{M}' : \Sigma \rightarrow \mathcal{L}(\mathcal{H}')$ such that $\mathbf{M}' \sim \mathbf{M}$ (i.e. $\mathbf{M}'(X) = 0$ exactly when $\mathbf{M}(X) = 0$) then $\mathbf{M}' \leq_{\text{pre}} \mathbf{M}$.

In the following subsections, we characterize the pre-processing maximal observables first in the case of discrete outcomes and then for the general case. The reason for this division is that we may formulate the maximality in a tighter fashion for discrete observables. Let us first recall a result from³¹:

Theorem 11. *Suppose that \mathcal{H} and \mathcal{H}' are separable Hilbert spaces, $\mathbf{P} : \Sigma \rightarrow \mathcal{L}(\mathcal{H}')$ is a sharp observable (a PVM), and $\mathbf{M} : \Sigma \rightarrow \mathcal{L}(\mathcal{H})$ is some observable such that $\mathbf{M} \ll \mathbf{P}$. There exists a channel $\Phi' : \mathcal{L}(\mathcal{H}') \rightarrow \mathcal{L}(\mathcal{H})$ such that $\mathbf{M}(X) = \Phi'(\mathbf{P}(X))$ for all $X \in \Sigma$, i.e., $\mathbf{M} \leq_{\text{pre}} \mathbf{P}$.*

A. Case of discrete observables

The theorem below characterizes pre-processing clean discrete observables.

Theorem 12. *Suppose that Ω is a finite or a countably infinite set and $\mathbf{M} : 2^\Omega \rightarrow \mathcal{L}(\mathcal{H})$ is an observable in a separable Hilbert space \mathcal{H} . Then \mathbf{M} is pre-processing clean if and only if it has the eigenvalue-1 property.*

Proof. We give the proof for the ‘only if’ part for an observable with a more general value space (Ω, Σ) where Σ is countably generated, since this yields no extra complications. Assume that $\mathbf{M} \in \text{Obs}(\Sigma, \mathcal{H})$ is pre-processing clean and μ is a probability measure on (Ω, Σ) such that $\mathbf{M} \sim \mu$. Also define the PVM $\mathbf{P}_\mu : \Sigma \rightarrow \mathcal{L}(L^2(\mu))$, $(\mathbf{P}_\mu(X)\psi)(x) = \chi_X(x)\psi(x)$ for all $X \in \Sigma$, all $\psi \in L^2(\mu)$, and μ -a.a. $x \in \Omega$. Hence, since $L^2(\mu)$ is separable, according to Theorem 11, $\mathbf{P}_\mu(X) = \Phi(\mathbf{M}(X))$ where $\Phi : \mathcal{L}(\mathcal{H}) \rightarrow \mathcal{L}(L^2(\mu))$ is a channel and thus, for all $\rho \in \mathcal{S}(L^2(\mu))$, $X \in \Sigma$, one has $\text{tr}[\rho\mathbf{P}_\mu(X)] = \text{tr}[\Phi_*(\rho)\mathbf{M}(X)]$. We define $\rho_X := \mu(X)^{-2}|\chi_X\rangle\langle\chi_X| \in \mathcal{S}(L^2(\mu))$ when $\mu(X) > 0$ (or $\mathbf{M}(X) \neq 0$). Let $\Phi_*(\rho_X) = \sum_{n=1}^r \lambda_n |\varphi_n\rangle\langle\varphi_n|$, $\lambda_n > 0$, $\sum_{n=1}^r \lambda_n = 1$, $\langle\varphi_n|\varphi_m\rangle = \delta_{nm}$, be the spectral decomposition of the state $\Phi_*(\rho_X)$. Now $1 = \text{tr}[\rho_X\mathbf{P}_\mu(X)] = \text{tr}[\Phi_*(\rho_X)\mathbf{M}(X)] = \sum_{n=1}^r \lambda_n \langle\varphi_n|\mathbf{M}(X)\varphi_n\rangle$ implying that $\langle\varphi_n|\mathbf{M}(X)\varphi_n\rangle = 1$ and thus $\|\sqrt{\mathbf{M}(\Omega \setminus X)}\varphi_n\|^2 = \langle\varphi_n|\mathbf{M}(\Omega \setminus X)\varphi_n\rangle = \langle\varphi_n|[I_{\mathcal{H}} - \mathbf{M}(X)]\varphi_n\rangle = 0$ or $\mathbf{M}(\Omega \setminus X)\varphi_n = \sqrt{\mathbf{M}(\Omega \setminus X)}\sqrt{\mathbf{M}(\Omega \setminus X)}\varphi_n = 0$ or $\mathbf{M}(X)\varphi_n = 1 \cdot \varphi_n$ for all $n < r + 1$.

Let us prove the ‘if’ part for a discrete observable \mathbf{M} which we identify with the effects $\mathbf{M}_i = \mathbf{M}(\{x_i\}) \neq 0$ such that $\sum_{i=1}^N \mathbf{M}_i = I_{\mathcal{H}}$, $N \in \mathbb{N}_\infty$ (without restricting generality, we assume that $\Omega = \{x_i\}_{i=1}^N$). Suppose then that any $\mathbf{M}(X)$, $X \neq \emptyset$, has the eigenvalue 1; denote the projection onto the corresponding eigenspace by P_X so that $P_X\mathbf{M}(X) = P_X = \mathbf{M}(X)P_X$. If $\emptyset \neq X \subseteq Y$ then $\langle\psi|\mathbf{M}(X)\psi\rangle \leq \langle\psi|\mathbf{M}(Y)\psi\rangle$ for all $\psi \in \mathcal{H}$ and $\mathbf{M}(X)\varphi = \varphi$ implies $\mathbf{M}(Y)\varphi = \varphi$, that is, $P_X \leq P_Y$. Similarly, if $X \cap Y = \emptyset$, then $P_X P_Y = 0$. Let us pick, for all $i < N + 1$, $\varphi_i \in P_{\{x_i\}}\mathcal{H}$, $\|\varphi_i\| = 1$, and define the projection $R := \sum_{i=1}^N |\varphi_i\rangle\langle\varphi_i|$. Using the above results, one immediately sees that $R\mathbf{M}_i R = |\varphi_i\rangle\langle\varphi_i|$. Define the channel

$\Psi : \mathcal{L}(\mathcal{H}) \rightarrow \mathcal{L}(R\mathcal{H})$, $\Psi(A) = RAR$. Now Ψ pre-processes M into the sharp observable $P : 2^\Omega \rightarrow \mathcal{L}(R\mathcal{H})$ defined via $P(\{x_i\}) := |\varphi_i\rangle\langle\varphi_i|$. If M was a pre-processing of another observable M' on 2^Ω (automatically $M' \sim M \sim P$) then P would also be a pre-processing of M' . According to Theorem 11, $M' \leq_{\text{pre}} P$ and, thus, $M' \leq_{\text{pre}} M$. \square

Note that the first part of the proof above shows that, even in the case where the value space (Ω, Σ) of a pre-processing maximal observable M is countably generated, M necessarily possesses the eigenvalue-1 property. This tells us that pre-processing maximal observables (with countably generated value spaces) determine their values and are not informationally complete.

B. Case of general observables

For the characterization of pre-processing clean observables with more general (countably generated) value spaces, we need first some auxiliary results. The proof of the following lemma follows closely the one of⁵ (Lemma 8.1).

Lemma 2. *Let $A \in \mathcal{L}(\mathcal{H})$, $0 \leq A \leq I_{\mathcal{H}}$, and $R \in \mathcal{P}(\mathcal{H})$, where \mathcal{H} is a Hilbert space. If $RAR \in \mathcal{P}(\mathcal{H})$ then A and R commute.*

Proof. Suppose that RAR is a projection. We now have

$$0 \leq ((I - R)AR)^*((I - R)AR) = RA(I - R)AR \leq RAR - RARAR = 0$$

implying that $(I - R)AR = 0$, i.e., $AR = RAR$. We obtain

$$AR = RAR = (RAR)^* = (AR)^* = RA,$$

proving the claim. \square

Let \mathcal{H} and \mathcal{K} be Hilbert spaces and $\Phi : \mathcal{L}(\mathcal{H}) \rightarrow \mathcal{L}(\mathcal{K})$ a channel. There is a minimal projection $R \in \mathcal{P}(\mathcal{H})$ such that $\Phi(R) = I_{\mathcal{K}}$, i.e., if $Q \in \mathcal{P}(\mathcal{H})$ is such that $\Phi(Q) = I_{\mathcal{K}}$ and $Q \leq R$, then $Q = R$. Thus, R can be called the support of Φ (indeed it is uniquely defined by Φ). Moreover, $\Phi(A) = \Phi(RAR)$ for all $A \in \mathcal{L}(\mathcal{H})$ and, whenever $A \in \mathcal{L}(\mathcal{H})$ is an effect, $\Phi(A) = 0$ if and only if $RAR = 0^5$ (Section 10.8). We now easily obtain the following result.

Lemma 3. *Let \mathcal{H} and \mathcal{K} be Hilbert spaces and $\Phi : \mathcal{L}(\mathcal{H}) \rightarrow \mathcal{L}(\mathcal{K})$ a channel with the support projection R . If $\Phi(A)$ is a projection for some effect $A \in \mathcal{L}(\mathcal{H})$, then RAR is a projection, $RA = AR$, and $A = RAR + R^\perp AR^\perp$.*

Proof. Let $A \in \mathcal{L}(\mathcal{H})$, $0 \leq A \leq I_{\mathcal{H}}$ and assume that $\Phi(RAR)$ is a projection. Using the Schwartz inequality (applicable especially to unital completely positive maps),

$$\Phi(RAR) = \Phi(RAR)^2 \leq \Phi(RARAR)$$

implying $\Phi(RAR - RARAR) \leq 0$. Since $RAR \geq RARAR$, one finds that $\Phi(RAR - RARAR) = 0$. Since $RAR - RARAR = R(A - ARA)R$ and $0 \leq A - ARA \leq I_{\mathcal{H}}$, the properties of the support projection cited above imply that $RAR = RARAR$, i.e., $RAR \in \mathcal{P}(\mathcal{H})$. Lemma 2 yields $AR = RA$ so that $A = RAR + RAR^\perp + R^\perp AR + R^\perp AR^\perp = RAR + R^\perp AR^\perp$. \square

Theorem 13. *Suppose that the σ -algebra $\Sigma \subseteq 2^\Omega$ is countably generated and $\mathbf{M} : \Sigma \rightarrow \mathcal{L}(\mathcal{H})$ is an observable operating in a separable Hilbert space \mathcal{H} . Then \mathbf{M} is pre-processing clean if and only if there exist a closed subspace $\mathcal{M} \subseteq \mathcal{H}$, a PVM $\mathbf{E} : \Sigma \rightarrow \mathcal{L}(\mathcal{M})$, and a POVM $\mathbf{F} : \Sigma \rightarrow \mathcal{L}(\mathcal{M}^\perp)$ such that $\mathbf{M} \sim \mathbf{E}$ and*

$$\mathbf{M}(X) = \mathbf{E}(X) \oplus \mathbf{F}(X), \quad X \in \Sigma. \quad (5)$$

Proof. Suppose that \mathbf{M} is pre-processing clean. Then \mathbf{M} can be pre-processed with a channel $\Phi : \mathcal{L}(\mathcal{H}) \rightarrow \mathcal{L}(L^2(\mu))$ into the observable \mathbf{P}_μ of the first part of the proof of Theorem 12, where $\mu \sim \mathbf{M}$. We have by Lemma 3 that there is an $R \in \mathcal{P}(\mathcal{H})$ such that $\Phi(R) = I_{L^2(\mu)}$ and $\mathbf{M}(X) = R\mathbf{M}(X)R + R^\perp\mathbf{M}(X)R^\perp$, where $R\mathbf{M}(X)R$ is a projection, for all $X \in \Sigma$. Suppose that $R\mathbf{M}(X)R = 0$ for some $X \in \Sigma$. Now $\mathbf{P}_\mu(X) = \Phi(R\mathbf{M}(X)R) = 0$ implying that $\mu(X) = 0$. Hence $\mathbf{M} \ll R\mathbf{M}(\cdot)R$; the contrary is, of course, automatically satisfied. Thus, we may choose $\mathcal{M} = R\mathcal{H}$ and $\mathbf{E}(X) = R\mathbf{M}(X)R$ and $\mathbf{F}(X) = R^\perp\mathbf{M}(X)R^\perp$ for all $X \in \Sigma$.

Assume then that the decomposition of \mathbf{M} into \mathbf{E} and \mathbf{F} of the claim exists. Clearly, $\mathbf{E} \leq_{\text{pre}} \mathbf{M}$ (with a rank-1 channel defined by the projection from \mathcal{H} onto \mathcal{M}). If $\mathbf{M}' : \Sigma \rightarrow \mathcal{L}(\mathcal{H}')$ is an observable such that $\mathbf{M}' \sim \mathbf{M} \sim \mathbf{E}$ and $\mathbf{M} \leq_{\text{pre}} \mathbf{M}'$ then, according to Theorem 11, $\mathbf{M}' \leq_{\text{pre}} \mathbf{E} \leq_{\text{pre}} \mathbf{M}$. \square

As also seen in the proof of Theorem 12, for $\mathbf{M} \in \text{Obs}(\Sigma, \mathcal{H})$ to be pre-processing clean, all the nonzero effects $\mathbf{M}(X)$ must have the eigenvalue 1. This is clearly satisfied by the direct

sum form (5) since \mathbf{E} is a PVM and, whenever $\mathbf{E}(X) = 0$, $\mathbf{M}(X) = 0$ as well. However, the contrary is problematic: if all nonzero $\mathbf{M}(X)$ have the eigenvalue 1, does it follow that we have the decomposition (5) with a fixed subspace \mathcal{M} ? This would mean that Theorem 12 extends plainly to general observables. We leave this as an open question. Moreover, in the finite-dimensional case, norm-1 observables also have the eigenvalue-1 property, which for (discrete) POVMs implies post-processing maximality; recall that now norm-1 POVMs are discrete if their outcome spaces are regular enough. An important infinite-dimensional and continuous counter example is the canonical phase observable Φ_{can} which is of norm 1, but none of its effects $\Phi_{\text{can}}(\Theta) \neq I_{\mathcal{H}}$ has the eigenvalue 1. Especially, Φ_{can} is not pre-processing maximal.

A different analysis of pre-processing can result in remarkably different characterizations of pre-processing clean observables. For instance, the authors of³ concentrate on finite-outcome observables on a *fixed* finite-dimensional Hilbert space \mathcal{H} . In this setting, an N -valued observable \mathbf{M} in \mathcal{H} is *clean* if for any N -valued observable \mathbf{M}' in \mathcal{H} such that there exists a channel $\Phi : \mathcal{L}(\mathcal{H}) \rightarrow \mathcal{L}(\mathcal{H})$ with $\mathbf{M}_i = \Phi(\mathbf{M}'_i)$, $i = 1, \dots, N$, there also exists a channel $\Psi : \mathcal{L}(\mathcal{H}) \rightarrow \mathcal{L}(\mathcal{H})$ such that $\mathbf{M}'_i = \Psi(\mathbf{M}_i)$, $i = 1, \dots, N$. With this definition, the set of clean observables within the set of N -valued observables in \mathcal{H} are exactly those \mathbf{M} such that $\|\mathbf{M}_i\| = 1$ for all $i = 1, \dots, N$, in the case where $N \leq \dim \mathcal{H}$. However, now also the case $N > \dim \mathcal{H}$ is possible and, in general, any rank-1 observable is clean. The difference in the definition of post-processing and clean observables of³ and the corresponding definitions of this paper is that, in³ one is restricted to using a single system within which to carry out pre-processing whereas in our analysis one is free to use any systems for pre-processing (no limitation to dimensionality of the Hilbert space from which one pre-processes).

Remark 3. Norm-1 observables, and hence also pre-processing maximal observables as eigenvalue-1 observables, are an example of so-called regular observables⁵ (Section 11.3): An effect $E \in \mathcal{L}(\mathcal{H})$ is called *regular* if $E \not\leq I_{\mathcal{H}} - E$ and $I_{\mathcal{H}} - E \not\leq E$ or, equivalently, the spectrum of E extends both above and below $1/2$. For example, a rank-1 effect $p|\varphi\rangle\langle\varphi|$, $p \in (0, 1]$, $\varphi \in \mathcal{H}$, $\|\varphi\| = 1$, is regular if and only if $p > \frac{1}{2}$ (if $\dim \mathcal{H} > 1$). An observable $\mathbf{M} : \Sigma \rightarrow \mathcal{L}(\mathcal{H})$ is *regular* if $\mathbf{M}(X)$ is regular whenever $0 \neq \mathbf{M}(X) \neq I_{\mathcal{H}}$. There exist regular POVMs which are not of norm-1 (e.g. $\mathbf{M}_1 = \frac{1}{3}|1\rangle\langle 1| + \frac{2}{3}|2\rangle\langle 2|$, $\mathbf{M}_2 = \frac{2}{3}|1\rangle\langle 1| + \frac{1}{3}|2\rangle\langle 2|$ is regular but not norm-1).

It is simple to check that whenever $\mathbf{M} : \Sigma \rightarrow \mathcal{L}(\mathcal{H})$ is regular, its range $\text{ran } \mathbf{M}$ equipped

with the intersection $\wedge_{\text{ran } \mathbf{M}}$,

$$\mathbf{M}(X) \wedge_{\text{ran } \mathbf{M}} \mathbf{M}(Y) := \inf\{\mathbf{M}(Z) \mid \mathbf{M}(Z) \leq \mathbf{M}(X), \mathbf{M}(Y)\}, \quad X, Y \in \Sigma,$$

and the complementation $' : \mathbf{M}(X) \mapsto \mathbf{M}(X)' = I_{\mathcal{H}} - \mathbf{M}(X)$ is a Boolean algebra, i.e., especially $\mathbf{M}(X) \wedge_{\text{ran } \mathbf{M}} \mathbf{M}(X)' = 0$ for all $X \in \Sigma$. In fact, the converse is true as well¹⁴: if $\mathbf{M} : \Sigma \rightarrow \mathcal{L}(\mathcal{H})$ is an observable such that $(\text{ran } \mathbf{M}, \wedge_{\text{ran } \mathbf{M}}, ')$ described above is a Boolean algebra, then \mathbf{M} is regular. Hence, a regular POVM \mathbf{M} preserves the ‘classical’ Boolean logic between the Boolean algebras Σ and $\text{ran } \mathbf{M}$.

Whether a regular observable can be informationally complete remains to be seen. This is not possible in the finite-dimensional case. To see this, let us consider an N -valued observable $\mathbf{M} = (\mathbf{M}_i)_{i=1}^N$ in a d -dimensional ($d < \infty$) Hilbert space \mathcal{H} . Taking the trace on both sides of the equation $I_{\mathcal{H}} = \sum_{i=1}^N \mathbf{M}_i$ and assuming that \mathbf{M} is informationally complete and regular, one arrives at $d = \sum_{i=1}^N \text{tr}[\mathbf{M}_i] > N/2 \geq d^2/2$, where the first inequality follows from regularity and the second from informational completeness. This is possible only if $d = 1$.

The above no-go result can be alleviated by relaxing the requirement on informational completeness within the set of *all* states. Let us consider an example in the two-dimensional Hilbert space. Fix the Pauli matrices σ_x , σ_y , and σ_z which in the eigenbasis of σ_z take the form

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Also denote $\mathbf{b} \cdot \boldsymbol{\sigma} := b_x \sigma_x + b_y \sigma_y + b_z \sigma_z$ for any $\mathbf{b} = (b_x, b_y, b_z) \in \mathbb{R}^3$. We now define the three-valued observable $\mathbf{M} = (\mathbf{M}_1, \mathbf{M}_2, \mathbf{M}_3)$ by $\mathbf{M}_i = 3^{-1}(I + \mathbf{a}_i \cdot \boldsymbol{\sigma})$, $i = 1, 2, 3$, where

$$\mathbf{a}_1 = (1, 0, 0), \quad \mathbf{a}_2 = \left(-\frac{1}{2}, \frac{\sqrt{3}}{2}, 0\right), \quad \mathbf{a}_3 = \left(-\frac{1}{2}, -\frac{\sqrt{3}}{2}, 0\right).$$

It is easily checked that \mathbf{M} is an extreme rank-1 observable and the non-zero eigenvalue of \mathbf{M}_i is $2/3$ for each $i = 1, 2, 3$. Thus, \mathbf{M} is regular. Moreover, \mathbf{M} is informationally complete within the restricted set of states ρ such that $\text{tr}[\rho \sigma_z] = 0$. Thus, \mathbf{M} is ‘more informationally complete’ than any PVM can be in the two-dimensional case. Indeed, whenever $\mathbf{P} = (|d_1\rangle\langle d_1|, |d_2\rangle\langle d_2|)$ is a PVM, the maximal subset of states where \mathbf{P} is informationally complete is the convex hull of the states $|d_1\rangle\langle d_1|$ and $|d_2\rangle\langle d_2|$ parametrized by a single parameter $p \in [0, 1]$ whereas one needs two parameters for the set of states ρ for which $\text{tr}[\rho \sigma_z] = 0$.

X. CONCLUSIONS

In this paper we have identified some important optimality properties of a quantum observable represented mathematically as a POVM \mathbf{M} : Determination of the past (the pre-measurement state of the system), i.e., informational completeness; freedom from dependence on more informational measurements from which the output data of \mathbf{M} may be processed, i.e., post-processing maximality; freedom from interference of different measurement schemes, i.e., extremality; determination of values, i.e., whether for any outcome set X one can prepare the system in a state realizing a value from X with arbitrarily high accuracy; determination of the future, i.e., any measurement of \mathbf{M} also works as a state preparator; and freedom from quantum noise, i.e., pre-processing maximality. We have investigated these properties and generalized results known for discrete observables for more general observables. We have also found connections between these conditions: Pre-processing maximality and determination of future are equivalent and both are characterized by the rank-1 property of \mathbf{M} . Moreover, using a refinement procedure, one can replace an informationally complete (resp. extreme) \mathbf{M} with an informationally complete (resp. extreme) rank-1 POVM \mathbf{M}^1 . Determination of values is equivalent with the norm-1 property (i.e. $\|\mathbf{M}(X)\| = 1$ for all outcome sets X , $\mathbf{M}(X) \neq I_{\mathcal{H}}$) and using our characterizations, we immediately see that, when \mathbf{M} is preprocessing clean, it automatically defines its values.

We may conclude that there are two major lines of optimality for quantum observables: On one hand, an observable may be informationally complete, and, at the same time, such an observable may also be free from all kinds of classical noise, i.e., it may be extreme and post-processing clean simultaneously. On the other hand, an observable may define its values, i.e., have the norm-1 property; especially, the observable subtype may be pre-processing clean. However, we are not aware of any norm-1 informationally complete observable; informational completeness requires properties that are strongly opposed to properties found in typical norm-1 observables: unsharpness, noncommutativity, and nonlocalizability, namely the inability to prepare systems into states yielding particular outcomes in the measurement with certainty. Thus the two main optimality criteria, informational completeness (determination of past) and determination of values, at its strongest in eigenvalue-1 observables, appear as complementary properties of a quantum observable. However, an observable may be free from classical noise (extreme rank-1) as well as from quantum noise (pre-processing

maximality) simultaneously, in which case the observable is forced to be a rank-1 PVM, which automatically defines its values and the future of the quantum system (only measurements of such observables are preparative) but fails to determine the past of the system.

We have also discussed and reviewed results concerning joint and sequential measurements involving optimal observables. Especially, all the observables which are jointly measurable with a rank-1 observable M are smearings (post-processings) of M , and for a jointly measurable pair (M, M') of observables, where either one of the observables is extreme, there exists a unique joint observable giving M and M' as its margins.

ACKNOWLEDGMENTS

The authors would like to thank T. Heinosaari for his remarks on the manuscript and Y. Kuramochi for turning the authors' attention to Ref.². E. H. also acknowledges financial support from the Japan Society for the Promotion of Science (JSPS) as an overseas postdoctoral fellow at JSPS.

REFERENCES

- ¹W. B. Arveson, "Subalgebras of C^* -algebras," *Acta Math.* **123**, 141-224 (1969)
- ²R. Beukema, "Maximality of positive operator-valued measures", *Positivity* **10**, 17-37 (2006)
- ³F. Buscemi, M. Keyl, G. M. D'Ariano, P. Perinotti, and R. F. Werner, "Clean positive operator valued measures", *J. Math. Phys.* **46**, 082109 (2005)
- ⁴P. Busch and P. Lahti, "The determination of the past and the future of a physical system in quantum mechanics", *Found. Phys.* **19**, 633-678 (1989)
- ⁵P. Busch, P. Lahti, J.-P. Pellonpää, and K. Ylisen, *Quantum Measurement* (Springer, 2016)
- ⁶C. Carmeli, T. Heinosaari, J. Schultz, and A. Toigo, "Tasks and premises in quantum state determination", *J. Phys. A: Math. Theor.* **47**, 075302 (2014)
- ⁷G. Chiribella, G. M. D'Ariano, and D. Schlingemann, "How continuous quantum measurements in finite dimensions are actually discrete", *Phys. Rev. Lett.* **98**, 190403 (2007)

- ⁸H. Cycon and K.-E. Hellwig, “Conditional expectations in generalized probability theory”, *J. Math. Phys.* **18**, 1154-1161 (1977)
- ⁹G. M. D’Ariano, P. Lo Presti, and P. Perinotti, “Classical randomness in quantum measurements”, *J. Phys. A: Math. Gen.* **38**, 5979 (2005)
- ¹⁰E. B. Davies, *Quantum Theory of Open Systems* (Academic Press, London, 1976)
- ¹¹E. B. Davies and J. T. Lewis, “An operational approach to quantum probability”, *Commun. Math. Phys.* **17**, 239-260 (1970)
- ¹²J. Dixmier, *Von Neumann Algebras* (North-Holland, Amsterdam, 1981)
- ¹³S.V. Dorofeev and J. de Graaf, “Some maximality results for effect-valued measures”, *Indag. Math. (N. S.)* **8(3)**, 349-369 (1997)
- ¹⁴P. Dvurečenskij and S. Pulmannová, “Difference posets, effects, and quantum measurements”, *Int. J. Theor. Phys.* **33**, 819-850 (1994)
- ¹⁵E. Haapasalo, T. Heinosaari, and J.-P. Pellonpää, “Quantum measurements on finite dimensional systems: relabeling and mixing,” *Quantum Inf. Process.* **11**, 1751-1763 (2012)
- ¹⁶E. Haapasalo, T. Heinosaari, and J.-P. Pellonpää, “When do pieces determine the whole? Extreme marginals of a completely positive map”, *Rev. Math. Phys.* **26**, 1450002 (2014)
- ¹⁷E. Haapasalo and J.-P. Pellonpää, “Extreme covariant quantum observables in the case of an Abelian symmetry group and a transitive value space”, *J. Math. Phys.* **52**, 122102 (2011)
- ¹⁸T. Heinosaari and T. Miyadera, “Universality of sequential quantum measurements”, *Phys. Rev. A* **91**, 022110 (2015)
- ¹⁹T. Heinosaari and J.-P. Pellonpää, “The canonical phase is pure”, *Phys. Rev. A* **80**, 040101 (R) (2009)
- ²⁰T. Heinosaari and J.-P. Pellonpää, “Generalized coherent states and extremal positive operator valued measures”, *J. Phys. A: Math. Theor.* **45**, 244019 (2012)
- ²¹T. Heinosaari and M. M. Wolf, “Non-disturbing quantum measurements”, *J. Math. Phys.* **51**, 092201 (2010)
- ²²A. S. Holevo, M. E. Shirokov, and R. F. Werner, “On the notion of entanglement in Hilbert spaces”, *Russian Math. Surveys* **60**, 359-360 (2005)
- ²³T. Hytönen, J.-P. Pellonpää, and K. Ylinen, “Positive sesquilinear form measures and generalized eigenvalue expansions,” *J. Math. Anal. Appl.* **336**, 1287-1304 (2007)
- ²⁴A. Jenčová, S. Pulmannová, and E. Vincecová, “Sharp and fuzzy observables on effect

- algebras”, *Int. J. Theor. Phys.* **47**, 125-148 (2008)
- ²⁵J. Kiukas, P. Lahti, J. Schultz, and R. F. Werner, “Characterization of informational completeness for covariant phase space observables”, *J. Math. Phys.* **53**, 102103 (2012)
- ²⁶J. Kiukas, J.-P. Pellonpää, and J. Schultz, “State reconstruction formulas for the s -distributions and quadratures”, *Rep. Math. Phys.* **66**, 55-84 (2010)
- ²⁷Y. Kuramochi, “Minimal sufficient positive-operator valued measure on a separable Hilbert space”, *J. Math. Phys.* **56**, 102205 (2015)
- ²⁸P. Lahti and K. Ylinen, “Dilations of positive operator measures and bimeasures related to quantum mechanics”, *Math. Slovaca* **54**, 169-189 (2004)
- ²⁹H. Martens and W. M. de Muynck, “Nonideal quantum measurements”, *Found. Phys.* **20**, 255-281 (1990)
- ³⁰K. R. Parthasarathy, “Extremal decision rules in quantum hypothesis testing”, *Infinite Dimens. Anal.* **2**, 557-568 (1999)
- ³¹J.-P. Pellonpää, “Complete characterization of extreme quantum observables in infinite dimensions”, *J. Phys. A: Math. Theor.* **44**, 085304 (2011)
- ³²J.-P. Pellonpää, “Quantum instruments: II. Measurement theory”, *J. Phys. A: Math. Theor.* **46**, 025303 (2013)
- ³³J.-P. Pellonpää, “Complete quantum measurements break entanglement”, *Phys. Lett. A* **376**, 3495 (2012)
- ³⁴J.-P. Pellonpää, “Complete measurements of quantum observables”, *Found. Phys.* **44**, 71-90 (2014)
- ³⁵J.-P. Pellonpää, “On coexistence and joint measurability of rank-1 quantum observables”, *J. Phys. A: Math. Theor.* **47**, 052002 (2014)
- ³⁶J.-P. Pellonpää and J. Schultz, “Measuring the canonical phase with phase-space measurements”, *Phys. Rev. A* **88**, 012121 (2013)
- ³⁷C. Preston, “Some notes on standard Borel and related spaces” (2008). Available at <https://arxiv.org/abs/0809.3066>
- ³⁸K. Ylinen, “Positive operator bimeasures and a noncommutative generalization”, *Studia Math.* **118**, 157 (1996)