KURENAI
Kyoto University Research Information Repository

KYOTO UNIVERSITY

Title	Minimal sufficient positive-operator valued measure on a separable Hilbert space
Author(s)	Kuramochi, Yui
Citation	Journal of Mathematical Physics (2015), 56(10)
Issue Date	2015-10-22
URL	http://hdl.handle.net/2433/203027
Right	© 2015 AIP Publishing LLC. This article may be downloaded for personal use only. Any other use requires prior permission of the author and the American Institute of Physics.
Type	Journal Article
Textversion	publisher

Kyoto University

Minimal sufficient positive-operator valued measure on a separable Hilbert space

Yui Kuramochi

Articles you may be interested in

Discrete Wigner function by symmetric informationally complete positive operator valued measure
J. Math. Phys. **50**, 072106 (2009); 10.1063/1.3177286

Intertwining operators between different Hilbert spaces: Connection with frames
J. Math. Phys. **50**, 043509 (2009); 10.1063/1.3094758

κ -Minkowski representations on Hilbert spaces
J. Math. Phys. **48**, 052305 (2007); 10.1063/1.2738360

A note on the infimum problem of Hilbert space effects
J. Math. Phys. **47**, 102103 (2006); 10.1063/1.2358392

Clean positive operator valued measures
J. Math. Phys. **46**, 082109 (2005); 10.1063/1.2008996
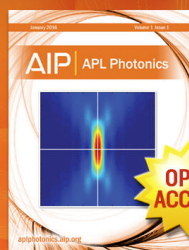
Minimal sufficient positive-operator valued measure on a separable Hilbert space

Yui Kuramochi[a)]
Department of Nuclear Engineering, Kyoto University, 6158540 Kyoto, Japan

We introduce a concept of a minimal sufficient positive-operator valued measure (POVM), which is the least redundant POVM among the POVMs that have the equivalent information about the measured quantum system. Assuming the system Hilbert space to be separable, we show that for a given POVM, a sufficient statistic called a Lehmann-Scheffé-Bahadur statistic induces a minimal sufficient POVM. We also show that every POVM has an equivalent minimal sufficient POVM and that such a minimal sufficient POVM is unique up to relabeling neglecting null sets. We apply these results to discrete POVMs and information conservation conditions proposed by the author. © *2015 AIP Publishing LLC.* [http://dx.doi.org/10.1063/1.4934235]

I. INTRODUCTION

A measurement outcome statistic of a general quantum measurement process is described by a positive-operator valued measure[1,2] (POVM). Such a description of quantum measurements enables us to formulate both projective and non-projective measurements. However, for the price of such a generality, some POVMs contain redundant information irrelevant to the system and, due to such redundancies, for a given POVM, there exist infinitely many POVMs that bring us the equivalent information about the system.

To clarify this point, let us consider the following example of a pair of discrete POVMs A and B:

$$A = \{A_0, A_1\},$$
$$O \leq A_0 \leq I, \, A_1 = I - A_0,$$
$$B = \{B_{00}, B_{01}, B_{10}, B_{11}\},$$
$$B_{i0} = \lambda A_i, \, B_{i1} = (1 - \lambda)A_i.$$

Here, $\lambda \in (0,1)$ and A_i and B_{ij} are bounded operators on a Hilbert space $\mathcal{H}$ corresponding to the measured quantum system and O and I are zero and identity operators, respectively. The measurement corresponding to the POVM B can be realized, for example, as follows: perform the measurement A, whose measurement outcome is $i \in \{0,1\}$, and generate a binary random variable j, which gives 0 with a probability λ and 1 with a probability $1 - \lambda$, and the measurement outcome of B is given by a pair (i, j). Apparently, $B = \{B_{ij}\}$ contains redundant information, which is in this case the classical random variable j, and A and B give the equivalent information about the system.

Then, it is natural to ask whether we can reduce such redundancies for a given POVM and how far such reductions proceed. In this paper, to formulate and answer this question, we introduce a concept of a minimal sufficient POVM which corresponds to the least redundant POVM among the POVMs that give the same information about the system. The main finding of this paper (Theorem 5) is that for any POVM on a separable Hilbert space, there exists a minimal sufficient POVM that has the equivalent information about the system and that such a minimal sufficient POVM is unique up to almost isomorphism, which is the relabeling neglecting null sets.

---

a)Electronic mail: kuramochi.yui.22c@st.kyoto-u.ac.jp

The concept of the minimal sufficient POVM has two origins: a minimal sufficient statistics[3,4] in mathematical statistics and a fuzzy equivalence relation[5–8] known in quantum measurement theory. Two POVMs are fuzzy equivalent if one of them can be realized by a classical post-processing of the other. The concept of the fuzzy equivalence relation is used in the definition of the minimal sufficient POVM. The minimal sufficiency condition for the POVM can be regarded as a generalization of the minimal sufficient statistics in the sense that we consider a more general class of post-processing which includes taking statistics.

This paper is organized as follows. In Sec. II, we show some preliminary results on mathematical statistics and quantum measurement theory which will be used in Sec. III. In Sec. III, we define two definitions of the minimal sufficient POVM corresponding to two kinds of fuzzy equivalence relations and introduce a sufficient statistic called a Lehmann-Scheffé-Bahadur (LSB) statistic for a given POVM. It is shown in Theorem 4 that the LSB statistic is a minimal sufficient statistic for the given POVM and a POVM induced by the LSB statistic is a minimal sufficient POVM. After introducing a concept of almost isomorphism between POVMs, we show in Theorem 5 that a POVM has an equivalent minimal sufficient POVM and that such a minimal sufficient POVM is unique up to almost isomorphism. In Sec. IV, we consider discrete POVMs and prove that for a given discrete POVM, there is a unique equivalent minimal sufficient POVM that is discrete and has no zero elements. In Sec. V, we apply the main results to information conservation conditions proposed by the author.[9,10]

II. PRELIMINARIES

In this section, we introduce some preliminary concepts and results on mathematical statistics and quantum measurement theory.

A. Sufficient and minimal sufficient statistics

Let $(\Omega, \mathcal{B})$ be a measurable space. A family of probability measures $\mathcal{P}$ with the outcome space $(\Omega, \mathcal{B})$ is called a statistical model on $(\Omega, \mathcal{B})$.

Let $\mathcal{P}$ be a statistical model on $(\Omega, \mathcal{B})$ and let λ be a probability measure on $(\Omega, \mathcal{B})$. $\mathcal{P}$ is said to be dominated by λ, denoted by $\mathcal{P} \ll \lambda$, if every element $P \in \mathcal{P}$ is absolutely continuous with respect to λ. A statistical model $\mathcal{P}$ is said to be dominated if there exists a probability measure that dominates $\mathcal{P}$. A dominated statistical model $\mathcal{P}$ has a countable subset $\{P_i\}_{i \geq 1} \subset \mathcal{P}$ such that $\lambda := \sum_{i \geq 1} c_i P_i$ dominates $\mathcal{P}$ if $c_i > 0$ and $\sum_{i \geq 1} c_i = 1$ hold.[11] Such λ is called a pivotal measure for $\mathcal{P}$.

Let $\mathcal{P} = \{P_\theta\}_{\theta \in \Theta}$ be a statistical model on an outcome space $(\Omega, \mathcal{B})$ and let $(\Omega_T, \mathcal{B}_T)$ be a measurable space. A $\mathcal{B}/\mathcal{B}_T$-measurable map $T \colon \Omega \to \Omega_T$ is called a *statistic*. The set of $\mathcal{B}/\mathcal{B}_T$-measurable maps (statistics) is denoted by $\mathbb{M}((\Omega, \mathcal{B}) \to (\Omega_T, \mathcal{B}_T))$. T is said to be *sufficient* with respect to the statistical model $\mathcal{P}$ if for every $E \in \mathcal{B}$, there exists a $\mathcal{B}_T$-measurable function $P(E|\cdot) \colon \Omega_T \to [0, \infty)$ such that $P_\theta(E|t) = P(E|t)P_\theta^T$-a.e. for each $P_\theta \in \mathcal{P}$, where $P_\theta(E|t)$ is the conditional probability of P_θ for given $T = t$ and $P_\theta^T(\cdot) := P_\theta(T^{-1}(\cdot))$.

Let P and Q be probability measures with an outcome space $(\Omega, \mathcal{B})$. Suppose $f \colon [0, \infty) \to \mathbb{R}$ be a strictly convex function such that $f(1) = 0$. Taking a σ-finite measure μ dominating P and Q, we write Radon-Nikodým derivatives as $p(x) := dP/d\mu(x)$ and $q(x) := dQ/d\mu(x)$. An f-divergence[12,13] between P and Q is defined by

$$D_f(P, Q) := \int_\Omega f\left(\frac{p(x)}{q(x)}\right) q(x) d\mu(x),$$

where the integral on the RHS is independent of the choice of μ and the following conventions are adopted:

$$f^*(0) := \lim_{t \to \infty} \frac{f(t)}{t},$$

$$0 \cdot f\left(\frac{p}{0}\right) = p f^*(0), \quad 0 \cdot f^*(0) = 0.$$

For later use, we fix an f such that $D_f(P,Q) < \infty$ for each P and Q. An example of such an f is given by $f(t) = (\sqrt{t} - 1)^2$, and the corresponding f-divergence is the Hellinger distance,

$$H(P,Q) = \int_\Omega \left(\sqrt{p(x)} - \sqrt{q(x)}\right)^2 d\mu(x) \le 2.$$

The following theorem gives necessary and sufficient conditions for the sufficiency of a statistic.

Theorem 1. *Let $\mathcal{P} = \{P_\theta\}_{\theta \in \Theta}$ be a statistical model on a measurable space $(\Omega, \mathcal{B})$ dominated by a pivotal probability measure λ, and let $T \colon \Omega \to \Omega_T$ be a $\mathcal{B}/\mathcal{B}_T$-measurable statistic and $P_\theta^T(\cdot) := P_\theta(T^{-1}(\cdot))$. Then, the following conditions are equivalent:*

(i) T is sufficient;

(ii)   for every $P_\theta \in \mathcal{P}$, there exists a $\mathcal{B}_T$-measurable real valued function $g_\theta(\cdot)$ such that

$$\frac{dP_\theta}{d\lambda}(x) = g_\theta(T(x)) \quad \lambda\text{-a.e.};$$

(iii)   for every $P_\theta \in \mathcal{P}$,

$$\frac{dP_\theta}{d\lambda}(x) = \frac{dP_\theta^T}{d\lambda^T}(T(x)) \quad \lambda\text{-a.e.},$$

where $\lambda^T(\cdot) := \lambda(T^{-1}(\cdot))$;

(iv) $D_f(P_{\theta_1}, P_{\theta_2}) = D_f(P_{\theta_1}^T, P_{\theta_2}^T)(\forall P_{\theta_1}, \forall P_{\theta_2} \in \mathcal{P})$.

Proof. The equivalence (i)⇔(ii)⇔(iv) is well-known.[4,11,13,14] The implication (iii)⇒(ii) is obvious. Let us show (ii)⇒(iii). For each $F \in \mathcal{B}_T$, we have

$$\begin{aligned}
P_\theta^T(F) &= \int_\Omega \chi_F(T(x)) \frac{dP_\theta}{d\lambda}(x) d\lambda(x) \\
&= \int_\Omega \chi_F(T(x)) g_\theta(T(x)) d\lambda(x) \\
&= \int_{\Omega_T} \chi_F(t) g_\theta(t) d\lambda^T(t),
\end{aligned}$$

where $\chi_F(\cdot)$ is the indicator function for F. This implies that $g_\theta(t) = dP_\theta^T / d\lambda^T(t) \lambda^T$-a.e. and we obtain (iii). □

Let $\mathcal{P}$ be a statistical model on $(\Omega, \mathcal{B})$. A statistic $T \in \mathbb{M}((\Omega, \mathcal{B}) \to (\Omega_T, \mathcal{B}_T))$ is said to be *minimal sufficient* if T is sufficient and for each sufficient statistic $S \in \mathbb{M}((\Omega, \mathcal{B}) \to (\Omega_S, \mathcal{B}_S))$, there exists a map $f \in \mathbb{M}((\Omega_S, \mathcal{B}_S) \to (\Omega_T, \mathcal{B}_T))$ such that $T(x) = f(S(x))\mathcal{P}$-a.e. A minimal sufficient statistic can be interpreted to capture the information about the statistical model $\mathcal{P}$ in the least redundant manner.

The following theorem gives a sufficient condition for the existence of a minimal sufficient statistic.

Theorem 2 (Lehmann and Scheffé[3] and Bahadur[4]). *Let $\mathcal{P} = \{P_\theta\}_{\theta \in \Theta}$ be a statistical model on an outcome space $(\Omega, \mathcal{B})$. Suppose that there exists a countable subset $\{P_{\theta_i}\}_{i \ge 1} \subset \mathcal{P}$ dense in $\mathcal{P}$ with respect to the following metric:*

$$d(P,Q) := \sup_{E \in \mathcal{B}} |P(E) - Q(E)|. \tag{1}$$

Then, a statistic $T \in \mathbb{M}((\Omega, \mathcal{B}) \to (\mathbb{R}^\infty, \mathcal{B}(\mathbb{R}^\infty)))$ defined by

$$T(x) := \left(\frac{dP_{\theta_i}}{d\lambda}(x)\right)_{i \ge 1} \in \mathbb{R}^\infty$$

is a minimal sufficient statistic. Here, $(\mathbb{R}^\infty, \mathcal{B}(\mathbb{R}^\infty))$ is the countable product space of the real line $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ and $\lambda := \sum_{i \ge 1} c_i P_{\theta_i}$ with $c_i > 0$ and $\sum_{i \ge 1} c_i = 1$.

B. Positive-operator valued measure

Throughout this paper, we fix a *separable* (i.e., $\dim \mathcal{H} \leq \aleph_0$) Hilbert space $\mathcal{H}$ and denote the set of bounded operators on $\mathcal{H}$ by $\mathcal{L}(\mathcal{H})$. A positive trace class operator ρ with unit trace is called a *state* on $\mathcal{H}$, and the set of states on $\mathcal{H}$ is denoted by $\mathcal{S}(\mathcal{H})$.

Let $(\Omega, \mathcal{B})$ be a measurable space. A POVM A is a mapping $A \colon \mathcal{B} \to \mathcal{L}(\mathcal{H})$ such that

(i) $A(E) \geq O$ for each $E \in \mathcal{B}$;
(ii) $A(\Omega) = I$;
(iii) for each countable disjoint family $\{E_i\} \subset \mathcal{B}$, $A(\bigcup_i E_i) = \sum_i A(E_i)$, where the RHS is convergent in the sense of the weak operator topology.

Here, O and I are the zero and identity operators, respectively. The triple $(\Omega, \mathcal{B}, A)$ is also called a POVM.

Let $(\Omega, \mathcal{B}, A)$ be a POVM. For each $\rho \in \mathcal{S}(\mathcal{H})$, we define a probability measure $P_\rho^A(\cdot)$ with the outcome space $(\Omega, \mathcal{B})$ by $P_\rho^A(E) := \mathrm{tr}[\rho A(E)]$ for each $E \in \mathcal{B}$. Then, A induces a natural statistical model $\mathcal{P}^A := \{P_\rho^A\}_{\rho \in \mathcal{S}(\mathcal{H})}$, which is the set of possible outcome distributions when we perform the measurement A.

From the separability of $\mathcal{H}$, there exists a sequence $\{\rho_i\}_{i \geq 1}$ in $\mathcal{S}(\mathcal{H})$ dense with respect to the trace norm topology. Taking arbitrary $\{c_i\}_{i \geq 1}$ such that $c_i > 0$ and $\sum_{i \geq 1} c_i = 1$, e.g., $c_i = 2^{-i}$, we define a state $\rho_* := \sum_{i \geq 1} c_i \rho_i$. Throughout this paper, we fix such $\{\rho_i\}_{i \geq 1}$ and ρ_*.

From the definition of ρ_*, $P_{\rho_*}^A = \sum_{i \geq 1} c_i P_{\rho_i}^A$ and the following proposition immediately follows.

Proposition 1. Let $(\Omega, \mathcal{B}, A)$ be a POVM. Then, $\mathcal{P}^A \ll P_{\rho_*}^A$, i.e., $P_{\rho_*}^A$ is a pivotal measure for $\mathcal{P}^A$.

Due to the above proposition, the notions of A-a.e., $\mathcal{P}^A$-a.e., and $P_{\rho_*}^A$-a.e. coincide, and thus, we will use them interchangeably.

C. Fuzzy preorder and equivalence relations among POVMs

Let $(\Omega_1, \mathcal{B}_1)$ and $(\Omega_2, \mathcal{B}_2)$ be measurable spaces. A mapping $\kappa(\cdot|\cdot) \colon \mathcal{B}_1 \times \Omega_2 \to [0, 1]$ is called a *regular Markov kernel* if

(i) $\kappa(E|\cdot)$ is $\mathcal{B}_2$-measurable for each $E \in \mathcal{B}_1$;
(ii) $\kappa(\cdot|y)$ is a probability measure for each $y \in \Omega_2$.

Let $(\Omega_1, \mathcal{B}_1)$ be a measurable space and let $(\Omega_2, \mathcal{B}_2, A_2)$ be a POVM. An A_2-*weak Markov kernel* is a mapping $\kappa(\cdot|\cdot) \colon \mathcal{B}_1 \times \Omega_2 \to \mathbb{R}$ such that

(i) $\kappa(E|\cdot)$ is $\mathcal{B}_2$-measurable for each $E \in \mathcal{B}_1$;
(ii) $0 \leq \kappa(E|y) \leq 1 A_2$-a.e. for each $E \in \mathcal{B}_1$;
(iii) $\kappa(\Omega_1|y) = 1 A_2$-a.e. and $\kappa(\emptyset|y) = 0 A_2$-a.e.;
(iv) $\kappa(\cup_i E_i|y) = \sum_i \kappa(E_i|y) A_2$-a.e. for each countable and disjoint $\{E_i\} \subset \mathcal{B}_1$.

By using the concepts of the regular and weak Markov kernels, we introduce fuzzy preorder and equivalence relations as follows.[6–8]

Definition 1. Let $(\Omega_A, \mathcal{B}_A, A)$ and $(\Omega_B, \mathcal{B}_B, B)$ be POVMs.

(i) *If there exists a regular Markov kernel $\kappa(\cdot|\cdot) \colon \mathcal{B}_A \times \Omega_B \to [0, 1]$ such that*

$$A(E) = \int_{\Omega_B} \kappa(E|y) dB(y) \quad (E \in \mathcal{B}_A), \tag{2}$$

then we say that A is regularly fuzzier than B, denoted by $A \preceq_r B$.

(ii) *If there exists a B-weak Markov kernel $\kappa(\cdot|\cdot) \colon \mathcal{B}_A \times \Omega_B \to \mathbb{R}$ such that condition (2) holds, then we say that A is weakly fuzzier than B, denoted by $A \preceq_w B$.*

(iii) *If $A \preceq_r B$ and $B \preceq_r A$, then A and B are said to be regularly fuzzy equivalent, denoted by $A \simeq_r B$.*

(iv)   If $A \preceq_w B$ and $B \preceq_w A$, then A and B are said to be weakly fuzzy equivalent, denoted by $A \simeq_w B$.

Intuitively, the relations $A \preceq_r B$ and $A \preceq_w B$ mean that the measurement of A can be realized by a classical post-processing of the measurement of B. Apparently, $A \preceq_r B$ (respectively, $A \simeq_r B$) implies $A \preceq_w B$ (respectively, $A \simeq_w B$). It is known[6,7,15] that the regular and weak relations $\simeq_r$ and $\simeq_w$ (respectively, $\preceq_r$ and $\preceq_w$) are equivalence relations (respectively, preorder relations). Apparently, the regular relation $A \simeq_r B$ implies the weak relation $A \simeq_w B$, while the converse does not necessarily hold. See the Appendix for an explicit example of POVMs that are weakly fuzzy equivalent but not regularly fuzzy equivalent.

A standard Borel space[16] is a measurable space Borel isomorphic to a complete separable metric space. We call a POVM with a standard Borel outcome space a standard Borel POVM. By further assuming the standard Borel properties of POVMs, the weak relations imply the corresponding regular relations as in the following proposition.

Proposition 2 (Remark 4.1 of Ref. 8). Let A and B be POVMs.

1.   *If A is standard Borel, $A \preceq_w B \Leftrightarrow A \preceq_r B$.*
2.   *If A and B are standard Borel, $A \simeq_w B \Leftrightarrow A \simeq_r B$.*

In the formulation of the minimal sufficient POVM, we will use the regular and weak fuzzy equivalence relations.

D. Sufficient statistics for POVM

In this subsection, we consider the sufficiency condition of a statistic for a POVM. We also show Lemma 2 which will be used in Sec. III.

Let $(\Omega_A, \mathscr{B}_A, A)$ be a POVM and let $T \in \mathbb{M}((\Omega_A, \mathscr{B}_A) \to (\Omega_T, \mathscr{B}_T))$ be a statistic. We define a POVM A_T with the outcome space $(\Omega_T, \mathscr{B}_T)$ by $A_T(\cdot) := A(T^{-1}(\cdot))$. Since $A_T(\cdot)$ can be written as

$$A_T(F) = \int_{\Omega_A} \chi_F(T(x)) dA(x) \quad (E \in \mathscr{B}_T),$$

we have $A_T \preceq_r A$. T is said to be a sufficient statistic for A if T is sufficient for the statistical model $\mathcal{P}^A = \{P_\rho^A\}_{\rho \in \mathcal{S}(\mathcal{H})}$.

The following lemma, the monotonicity of the f-divergence, states that the f-divergence is monotonically decreasing by the classical post-processing.

Lemma 1 (Theorem 7.1 of Ref. 8). Let A and B be POVMs.

(i) $A \preceq_w B$ implies $D_f(P_\rho^A, P_\sigma^A) \leq D_f(P_\rho^B, P_\sigma^B)(\forall \rho, \forall \sigma \in \mathcal{S}(\mathcal{H}))$.
(ii) $A \simeq_w B$ implies $D_f(P_\rho^A, P_\sigma^A) = D_f(P_\rho^B, P_\sigma^B)(\forall \rho, \forall \sigma \in \mathcal{S}(\mathcal{H}))$.

The next theorem characterizes the sufficiency of a statistic for a POVM.

Theorem 3. *Let $(\Omega_A, \mathscr{B}_A, A)$ be a POVM and let $T \in \mathbb{M}((\Omega_A, \mathscr{B}_A) \to (\Omega_T, \mathscr{B}_T))$ be a statistic. Then, the following conditions are equivalent:*

(i)   *T is sufficient for A;*
(ii)   *for each $\rho \in \mathcal{S}(\mathcal{H})$, there exists a $\mathscr{B}_T$-measurable map $g_\rho(\cdot)$ such that*

$$\frac{dP_\rho^A}{dP_{\rho_*}^A}(x) = g_\rho(T(x)) \quad A\text{-a.e.};$$

(iii)   *for each $\rho \in \mathcal{S}(\mathcal{H})$,*

$$\frac{dP_\rho^A}{dP_{\rho_*}^A}(x) = \frac{dP_\rho^{A_T}}{dP_{\rho_*}^{A_T}}(T(x)) \quad A\text{-a.e.};$$

(iv)   *for each $\rho, \sigma \in \mathcal{S}(\mathcal{H})$, $D_f(P_\rho^A, P_\sigma^A) = D_f(P_\rho^{A_T}, P_\sigma^{A_T})$;*

(v) $A \simeq_w A_T$.

Furthermore, if A is a standard Borel POVM the above conditions are equivalent to

(vi) $A \simeq_r A_T$.

Proof. From Theorem 1 and Lemma 1, the equivalence (i)⇔(ii)⇔(iii)⇔(iv) immediately follows.

Let us show (i)⇒(v). $A_T \preceq_r A$ is evident from the definition of A_T. Since T is sufficient, there exists a conditional probability $P(E|\cdot)$, which is $\mathscr{B}_T$-measurable for each $E \in \mathscr{B}_A$, such that

$$P_\rho^A(E) = \int_{\Omega_T} P(E|t) \, dP_\rho^{A_T}(t).$$

Since $P(\cdot|\cdot)$ satisfies the conditions for the A_T-weak Markov kernel from the definition of the conditional probability, we obtain $A \preceq_w A_T$, and thus, (v) holds.

(v)⇒(iv) follows from Lemma 1.

Let us assume A is a standard Borel POVM. If we assume (v), from Proposition 2, we have $A \preceq_r A_T$. Since $A_T \preceq_r A$ by definition, we have proved (v)⇒(vi). The converse (vi)⇒(v) is evident.   □

The following lemma assures the existence of a POVM corresponding to the joint distribution for a given POVM and a regular Markov kernel.

Lemma 2. Let $(\Omega_1, \mathscr{B}_1)$ be a measurable space and let $(\Omega_2, \mathscr{B}_2, A_2)$ be a POVM. Let $\kappa(\cdot|\cdot) \colon \mathscr{B}_1 \times \Omega_2 \to [0, 1]$ be a regular Markov kernel. Then, the following assertions hold.

1.   There exists a unique POVM A_{12} with the product outcome space $(\Omega_1 \times \Omega_2, \mathscr{B}_1 \times \mathscr{B}_2)$ such that

$$A_{12}(E_1 \times E_2) = \int_{E_2} \kappa(E_1|y) dA_2(y) \quad (E_1 \in \mathscr{B}_1, \, E_2 \in \mathscr{B}_2). \tag{3}$$

2.   The canonical projection

$$\pi_2 \colon \Omega_1 \times \Omega_2 \ni (x, y) \mapsto y \in \Omega_2$$

is a sufficient statistic for A_{12}.

3. $A_2 = (A_{12})_{\pi_2} \simeq_r A_{12}$.

4.   For each $\rho \in \mathcal{S}(\mathcal{H})$,

$$\frac{dP_\rho^{A_{12}}}{dP_{\rho_*}^{A_{12}}}(x, y) = \frac{dP_\rho^{A_2}}{dP_{\rho_*}^{A_2}}(y) \quad A_{12}\text{-a.e.}$$

Proof. 1.   Let us consider the following mapping:

$$\tilde{\kappa}(F|y) := \kappa(F|_y|y), \quad (F \in \mathscr{B}_1 \times \mathscr{B}_2, y \in \Omega_2),$$
$$F|_y := \{x \in \Omega_1 | (x, y) \in F\}.$$

Then, $\tilde{\kappa}(\cdot|y)$ is a probability measure for each $y \in \Omega_2$. To show the measurability of $\tilde{\kappa}(F|\cdot)$ for each $F \in \mathscr{B}_1 \times \mathscr{B}_2$, let us define a class

$$\mathscr{D}_{\tilde{\kappa}} := \{F \in \mathscr{B}_1 \times \mathscr{B}_2 | \tilde{\kappa}(F|\cdot) \text{ is } \mathscr{B}_2\text{–measurable}\}.$$

Then, $\mathscr{D}_{\tilde{\kappa}}$ is a Dynkin class, i.e., $\mathscr{D}_{\tilde{\kappa}}$ contains $\Omega_1 \times \Omega_2$ and is closed under proper differences and countable disjoint unions. Since

$$\tilde{\kappa}(E_1 \times E_2|y) = \kappa(E_1|y)\chi_{E_2}(y) \quad (E_1 \in \mathscr{B}_1, E_2 \in \mathscr{B}_2), \tag{4}$$

$\mathscr{D}_{\tilde{\kappa}}$ contains the family of cylinder sets $\mathscr{C} := \{E_1 \times E_2 \,|\, E_1 \in \mathscr{B}_1, E_2 \in \mathscr{B}_2\}$, which is a multiplicative class. Then, the Dynkin class theorem assures that $\mathscr{D}_{\tilde{\kappa}} = \mathscr{B}_1 \times \mathscr{B}_2$, i.e., $\tilde{\kappa}(\cdot|\cdot)$ is a regular Markov kernel. Let us define a POVM A_{12} by

$$A_{12}(F) := \int_{\Omega_2} \tilde{\kappa}(F|y) dA_2(y), \quad (F \in \mathscr{B}_1 \times \mathscr{B}_2). \tag{5}$$

Then from Eq. (4), A_{12} satisfies desired condition (3). The uniqueness of A_{12} immediately follows from the Dynkin class theorem.

2. Let us denote the conditional probability of $P_\rho^{A_{12}}$ for given $\pi_2 = y$ by $P_\rho^{A_{12}}(F|y)(F \in \mathscr{B}_1 \times \mathscr{B}_2)$. Then for each $E_1 \in \mathscr{B}_1$ and $E_2, E_2' \in \mathscr{B}_2$, we have

$$
\begin{aligned}
P_\rho^{A_{12}}(E_1 \times (E_2 \cap E_2')) &= \int_{E_2'} \kappa(E_1|y)\chi_{E_2}(y)dP_\rho^{A_2}(y) \\
&= \int_{E_2'} P_\rho^{A_{12}}(E_1 \times E_2|y)dP_\rho^{A_2}(y).
\end{aligned}
$$

Thus, we have

$$
P_\rho^{A_{12}}(E_1 \times E_2|y) = \kappa(E_1|y)\chi_{E_2}(y) \quad P_\rho^{A_2}\text{-a.e.}
$$

and

$$
P_{\rho_*}^{A_{12}}(E_1 \times E_2|y) = \kappa(E_1|y)\chi_{E_2}(y) \quad A_2\text{-a.e.}
$$

Therefore, we obtain

$$
P_\rho^{A_{12}}(E_1 \times E_2|y) = P_{\rho_*}^{A_{12}}(E_1 \times E_2|y) \quad P_\rho^{A_2}\text{-a.e.}
$$

To show that $P_\rho^{A_{12}}(F|y)$ can be taken independent of ρ, we define a class

$$
\mathscr{D}_\rho := \{F \in \mathscr{B}_1 \times \mathscr{B}_2 \,|\, P_\rho^{A_{12}}(F|y) = P_{\rho_*}^{A_{12}}(F|y) \quad P_\rho^{A_2}\text{-a.e.}\}.
$$

Then, $\mathscr{D}_\rho$ is a Dynkin class that contains the family of cylinder sets $\mathscr{C}$. Therefore, the Dynkin class theorem assures that $P_\rho^{A_{12}}(F|y) = P_{\rho_*}^{A_{12}}(F|y)P_\rho^{A_2}$-a.e. for each $F \in \mathscr{B}_1 \times \mathscr{B}_2$. This implies that π_2 is a sufficient statistic.

The assertion 3 is evident from Eq. (5) and $A_2 = (A_{12})_{\pi_2} \preceq_r A_{12}$.

The assertion 4 immediately follows from assertion 3 and Theorem 3.   □

III. MINIMAL SUFFICIENT POVM

In this section, we define a minimal sufficient POVM and show the existence and uniqueness of a minimal sufficient POVM equivalent to a given POVM.

Definition 2. Let $(\Omega_A, \mathscr{B}_A, A)$ be a POVM.

1. *$(\Omega_A, \mathscr{B}_A, A)$ is said to be $\simeq_r$-minimal sufficient if for any POVM $(\Omega_B, \mathscr{B}_B, B)$ regularly fuzzy equivalent to A, there exists a measurable map $f \in \mathbb{M}((\Omega_B, \mathscr{B}_B) \to (\Omega_A, \mathscr{B}_A))$ such that $B_f = A$.*
2. *$(\Omega_A, \mathscr{B}_A, A)$ is said to be $\simeq_w$-minimal sufficient if for any POVM $(\Omega_B, \mathscr{B}_B, B)$ weakly fuzzy equivalent to A, there exists a measurable map $f \in \mathbb{M}((\Omega_B, \mathscr{B}_B) \to (\Omega_A, \mathscr{B}_A))$ such that $B_f = A$.*

Since $\simeq_w$ is a relation less restrictive than $\simeq_r$, any $\simeq_w$-minimal sufficient POVM is a $\simeq_r$-minimal sufficient POVM by definition.

We remark that the measurable map f in Definition 2 is a sufficient statistic for the statistical model $\mathcal{P}^B$ due to Theorem 3 and $B_f = A \simeq_w B$.

The minimal sufficiency of a POVM can be interpreted as a generalization of the minimal sufficiency of a statistic in the sense that we consider a more general class of classical post-processings which includes taking statistics. Intuitively, a minimal sufficient POVM A is the least redundant POVM among POVMs fuzzy equivalent to A.

A. Lehmann-Scheffé-Bahadur statistic

Let $(\Omega_A, \mathscr{B}_A, A)$ be a POVM. Corresponding to Theorem 2, we define a statistic $T \in \mathbb{M}((\Omega_A, \mathscr{B}_A) \to (\mathbb{R}^\infty, \mathscr{B}(\mathbb{R}^\infty)))$ by

$$
T(x) := \left(\frac{dP_{\rho_i}^A}{dP_{\rho_*}^A}(x)\right)_{i \geq 1} \in \mathbb{R}^\infty. \tag{6}
$$

By taking authors' names of Refs. 3 and 4, we call statistic (6) a *LSB statistic* for A.

The following theorem is the first main result of this paper.

Theorem 4. *Let $(\Omega_A, \mathscr{B}_A, A)$ be an arbitrary POVM and let T be the LSB statistic given by (6). Then, the following assertions hold.*

1. *T is a minimal sufficient statistic for $\mathcal{P}^A$.*
2. *A_T is a $\simeq_w$-minimal sufficient POVM and therefore a $\simeq_r$-minimal sufficient POVM.*

Proof. 1.   From Theorem 2, it is sufficient to show that $\{P^A_{\rho_i}\}_{i \geq 1}$ is dense in $\mathcal{P}^A$ with respect to the metric $d(\cdot, \cdot)$ given by (1). For each $\rho, \sigma \in \mathcal{S}(\mathcal{H})$ and each $E \in \mathscr{B}_A$, we have

$$|P^A_\rho(E) - P^A_\sigma(E)| = |\mathrm{tr}[(\rho - \sigma)A(E)]|$$
$$\leq \|\rho - \sigma\|_{\mathrm{tr}}\|A(E)\| \leq \|\rho - \sigma\|_{\mathrm{tr}},$$

where $\| \cdot \|$ and $\| \cdot \|_{\mathrm{tr}}$ are the operator and trace norms, respectively. Thus, we obtain

$$d(P^A_\rho, P^A_\sigma) \leq \|\rho - \sigma\|_{\mathrm{tr}}. \tag{7}$$

Since $\{\rho_i\}_{i \geq 1}$ is dense in $\mathcal{S}(\mathcal{H})$ with respect to the trace norm, inequality (7) implies that $\{P^A_{\rho_i}\}_{i \geq 1}$ is dense with respect to $d(\cdot, \cdot)$ in $\mathcal{P}^A$, and we have shown the minimal sufficiency of T.

2.   From Theorem 3 and the sufficiency of T, we have

$$\frac{dP^A_{\rho_i}}{dP^A_{\rho_*}}(x) = \frac{dP^{A_T}_{\rho_i}}{dP^{A_T}_{\rho_*}}(T(x)) \quad A\text{-a.e.}$$

for each $i \geq 1$. From definition of the LSB statistic (6), this implies

$$T(x) = \left(\frac{dP^{A_T}_{\rho_i}}{dP^{A_T}_{\rho_*}}(T(x)) \right)_{i \geq 1} \quad A\text{-a.e.}$$

or

$$t = \left(\frac{dP^{A_T}_{\rho_i}}{dP^{A_T}_{\rho_*}}(t) \right)_{i \geq 1} \quad A_T\text{-a.e.} \tag{8}$$

To show the $\simeq_w$-minimal sufficiency of A_T, let $(\Omega_B, \mathscr{B}_B, B)$ be an arbitrary POVM weakly fuzzy equivalent to A_T. Since the outcome space $(\mathbb{R}^\infty, \mathscr{B}(\mathbb{R}^\infty))$ for A_T is a standard Borel space, there exists a regular Markov kernel $\kappa(\cdot|\cdot)\colon \mathscr{B}(\mathbb{R}^\infty) \times \Omega_B \to [0,1]$ such that

$$A_T(F) = \int_{\Omega_B} \kappa(F|y)dB(y) \tag{9}$$

for each $F \in \mathscr{B}(\mathbb{R}^\infty)$. We define a POVM $(\mathbb{R}^\infty \times \Omega_B, \mathscr{B}(\mathbb{R}^\infty) \times \mathscr{B}_B, C)$ by

$$C(E) \coloneqq \int_{\Omega_B} \kappa(E|_y|y)dB(y) \quad (E \in \mathscr{B}(\mathbb{R}^\infty) \times \mathscr{B}_B).$$

Then from Lemma 2, we have $C \simeq_r B = C_{\pi_B}$ and

$$\frac{dP^C_{\rho_i}}{dP^C_{\rho_*}}(t, y) = \frac{dP^B_{\rho_i}}{dP^B_{\rho_*}}(y) \quad C\text{-a.e.,} \tag{10}$$

for each $i \geq 1$, where $\pi_B\colon \mathbb{R}^\infty \times \Omega_B \to \Omega_B$ is the canonical projection. On the other hand, from Eq. (9), we have $C_{\pi_{A_T}} = A_T \simeq_w B \simeq_r C$, where $\pi_{A_T}\colon \mathbb{R}^\infty \times \Omega_B \to \mathbb{R}^\infty$ is the canonical projection,

and Theorem 3 assures that

$$\frac{dP^C_{\rho_i}}{dP^C_{\rho_*}}(t, y) = \frac{dP^{A_T}_{\rho_i}}{dP^{A_T}_{\rho_*}}(t) \quad C\text{-a.e.} \tag{11}$$

for each $i \geq 1$. From Eqs. (8), (10) and (11), we obtain

$$t = \left(\frac{dP^B_{\rho_i}}{dP^B_{\rho_*}}(y)\right)_{i \geq 1} =: f(y) \quad C\text{-a.e.,} \tag{12}$$

where $f \in \mathbb{M}((\Omega_B, \mathcal{B}_B) \to (\mathbb{R}^\infty, \mathcal{B}(\mathbb{R}^\infty)))$. Then for each $E \in \mathcal{B}(\mathbb{R}^\infty)$, we have

$$
\begin{aligned}
B_f(E) &= B(f^{-1}(E)) \\
&= C(\mathbb{R}^\infty \times f^{-1}(E)) \\
&= \int_{\mathbb{R}^\infty \times \Omega_B} \chi_E(f(y)) dC(t, y) \\
&= \int_{\mathbb{R}^\infty \times \Omega_B} \chi_E(t) dC(t, y) \\
&= A_T(E),
\end{aligned}
\tag{13}
$$

where we have used Eq. (12) in deriving equality (13). Thus, we obtain $B_f = A_T$, and we have shown the $\simeq_w$-minimal sufficiency of A_T. □

B. Uniqueness up to almost isomorphism

In order to formulate the uniqueness of the minimal sufficient POVM, we introduce a concept of almost isomorphism as follows.

Definition 3. Let $(\Omega_i, \mathcal{B}_i, A_i)(i = 1, 2)$ be POVMs.

(i) *A $\mathcal{B}_1/\mathcal{B}_2$-bimeasurable bijection $f : \Omega_1 \to \Omega_2$ is called a strict isomorphism if $(A_1)_f = A_2$. If there exists such a strict isomorphism, $(\Omega_1, \mathcal{B}_1, A_1)$ and $(\Omega_2, \mathcal{B}_2, A_2)$ are said to be strictly isomorphic, denoted by $(\Omega_1, \mathcal{B}_1, A_1) \approx (\Omega_2, \mathcal{B}_2, A_2)$.*

(ii) *$(\Omega_1, \mathcal{B}_1, A_1)$ and $(\Omega_2, \mathcal{B}_2, A_2)$ are said to be almost isomorphic, written as $(\Omega_1, \mathcal{B}_1, A_1) \sim (\Omega_2, \mathcal{B}_2, A_2)$, if there exist measurable subsets $\tilde{\Omega}_i \in \mathcal{B}_i(i = 1, 2)$ such that $A_i(\tilde{\Omega}_i) = I$ and the restrictions $(\tilde{\Omega}_i, \tilde{\Omega}_i \cap \mathcal{B}_i, A_i|_{\tilde{\Omega}_i})(i = 1, 2)$ are strictly isomorphic. Here, $\tilde{\Omega}_i \cap \mathcal{B}_i := \{\tilde{\Omega}_i \cap E \mid E \in \mathcal{B}_i\}$ and $A_i|_{\tilde{\Omega}_i}$ is the restriction of A_i to $\tilde{\Omega}_i \cap \mathcal{B}_i$. We call a measurable subset $\tilde{\Omega}_i$ with $A_i(\tilde{\Omega}_i) = I$ a full measure set.*

The above definition is a straightforward generalization of the corresponding concepts known in classical measures.[17,18] The relations $\approx$ and $\sim$ are equivalence relations, which can be proved in a similar manner as for the classical measure (e.g., Sec. 2.4 of Ref. 17). Intuitively, these concepts correspond to the relabeling of the measurement outcomes.

The following proposition gives the relationship between these isomorphisms and the fuzzy equivalence relation.

Proposition 3. Let $(\Omega_A, \mathcal{B}_A, A)$ and $(\Omega_B, \mathcal{B}_B, B)$ be POVMs. Then, the following implications hold.

$$(\Omega_A, \mathcal{B}_A, A) \approx (\Omega_B, \mathcal{B}_B, B) \Rightarrow (\Omega_A, \mathcal{B}_A, A) \sim (\Omega_B, \mathcal{B}_B, B)$$

$$\Rightarrow A \simeq_r B.$$

Proof. The first implication is evident from the definitions of the strict and almost isomorphisms. Let us assume $(\Omega_A, \mathcal{B}_A, A) \sim (\Omega_B, \mathcal{B}_B, B)$ and show $A \simeq_r B$. Let $\tilde{\Omega}_A \in \mathcal{B}_A$ and $\tilde{\Omega}_B \in \mathcal{B}_B$ be full measure sets such that $(\tilde{\Omega}_A, \tilde{\Omega}_A \cap \mathcal{B}_A, A|_{\tilde{\Omega}_A}) \approx (\tilde{\Omega}_B, \tilde{\Omega}_B \cap \mathcal{B}_B, B|_{\tilde{\Omega}_B})$. We first show $A \simeq_r A|_{\tilde{\Omega}_A}$. Since the

identity map

$$\iota \colon \tilde{\Omega}_A \ni x \mapsto x \in \Omega_A$$

is $\tilde{\Omega}_A \cap \mathscr{B}_A / \mathscr{B}_A$-measurable, we have $A = (A|_{\tilde{\Omega}_A})_\iota \preceq_r A|_{\tilde{\Omega}_A}$. We fix a point $x_0 \in \tilde{\Omega}_A$ and define a mapping $\tilde{\iota} \colon \Omega_A \to \tilde{\Omega}_A$ by

$$\tilde{\iota}(x) := \begin{cases} x & (x \in \tilde{\Omega}_A) \\ x_0 & (x \in \Omega_A \setminus \tilde{\Omega}_A) \end{cases}. \tag{14}$$

Then, for each $E \in \tilde{\Omega}_A \cap \mathscr{B}_A$, we have

$$\tilde{\iota}^{-1}(E) = \begin{cases} E \cup (\Omega_A \setminus \tilde{\Omega}_A) & (x_0 \in E) \\ E & (x_0 \notin E) \end{cases}.$$

Hence, $\tilde{\iota}$ is $\mathscr{B}_A / \tilde{\Omega}_A \cap \mathscr{B}_A$-measurable and $A|_{\tilde{\Omega}_A} = A_{\tilde{\iota}} \preceq_r A$. Thus, we have shown $A \simeq_r A|_{\tilde{\Omega}_A}$.

Similarly we can prove $B \simeq_r B|_{\tilde{\Omega}_B}$. Since the equivalence $A|_{\tilde{\Omega}_A} \simeq_r B|_{\tilde{\Omega}_B}$ immediately follows from the definition of the strict isomorphism, we obtain $A \simeq_r A|_{\tilde{\Omega}_A} \simeq_r B|_{\tilde{\Omega}_B} \simeq_r B$. $\qquad\square$

We next show that the minimal sufficiency condition for a POVM is invariant under almost isomorphism.

Proposition 4. Let $(\Omega_A, \mathscr{B}_A, A)$ be a $\simeq_w$-minimal sufficient (respectively, $\simeq_r$-minimal sufficient) POVM and let $(\Omega_B, \mathscr{B}_B, B)$ be a POVM almost isomorphic to $(\Omega_A, \mathscr{B}_A, A)$. Then, B is also $\simeq_w$-minimal sufficient (respectively, $\simeq_r$-minimal sufficient).

Proof. We prove the assertion for the $\simeq_w$-minimal sufficient POVM; the proof for $\simeq_r$-minimal sufficiency can be obtained by replacing $\simeq_w$ to $\simeq_r$ in the following proof.

Let $\tilde{\Omega}_A \in \mathscr{B}_A$ and $\tilde{\Omega}_B \in \mathscr{B}_B$ be full measure sets such that $(\tilde{\Omega}_A, \tilde{\Omega}_A \cap \mathscr{B}_A, A|_{\tilde{\Omega}_A}) \approx (\tilde{\Omega}_B, \tilde{\Omega}_B \cap \mathscr{B}_B, B|_{\tilde{\Omega}_B})$ by a strict isomorphism $f \colon \tilde{\Omega}_A \to \tilde{\Omega}_B$. To show the $\simeq_w$-minimal sufficiency of B, we take an arbitrary POVM $(\Omega_C, \mathscr{B}_C, C)$ weakly fuzzy equivalent to B. Then from Proposition 3, C is also weakly fuzzy equivalent to A and there exists a mapping $g \in \mathbb{M}((\Omega_C, \mathscr{B}_C) \to (\Omega_A, \mathscr{B}_A))$ such that $C_g = A$. By fixing a point $x_0 \in \tilde{\Omega}_A$, we define a mapping $\tilde{\iota} \colon \Omega_A \to \tilde{\Omega}_A$ by Eq. (14). Then, $\tilde{\iota}$ is a measurable map such that $A|_{\tilde{\Omega}_A} = A_{\tilde{\iota}}$. Let us denote by $\tilde{f}$ the isomorphism f regarded as a map from $\tilde{\Omega}_A$ to Ω_B. Then, $\tilde{f}$ is $\tilde{\Omega}_A \cap \mathscr{B}_A / \mathscr{B}_B$-measurable and $(A|_{\tilde{\Omega}_A})_{\tilde{f}} = B$. Thus, we have $B = (A|_{\tilde{\Omega}_A})_{\tilde{f}} = (A_{\tilde{\iota}})_{\tilde{f}} = ((C_g)_{\tilde{\iota}})_{\tilde{f}} = C_{\tilde{f} \circ \tilde{\iota} \circ g}$, and B is $\simeq_w$-minimal sufficient. $\qquad\square$

Now we prove the following theorem which is the second main result of this paper.

Theorem 5. *1. Let $(\Omega_A, \mathscr{B}_A, A)$ be an arbitrary POVM. Then, there exists a $\simeq_w$-minimal sufficient POVM $\tilde{A}$ weakly fuzzy equivalent to A. Furthermore, such $\tilde{A}$ is unique up to almost isomorphism.*

2. Let $(\Omega_A, \mathscr{B}_A, A)$ be a standard Borel POVM. Then, there exists a $\simeq_r$-minimal sufficient POVM $\tilde{A}$ regularly fuzzy equivalent to A. Furthermore, such $\tilde{A}$ is unique up to almost isomorphism.

Proof. 1. Let $T \colon \Omega_A \to \mathbb{R}^\infty$ be the LSB statistic given by (6) and define $\tilde{A} := A_T$. Then, from Theorems 3 and 4, $\tilde{A}$ is a $\simeq_w$-minimal sufficient POVM weakly fuzzy equivalent to A. From the same discussion in Theorem 4, $\tilde{A}$ satisfies the following condition:

$$t = \left(\frac{dP_{\rho_i}^{\tilde{A}}}{dP_{\rho_*}^{\tilde{A}}}(t) \right)_{i \geq 1} \qquad \tilde{A}\text{-a.e.} \tag{15}$$

To prove the uniqueness up to almost isomorphism, take another $\simeq_w$-minimal sufficient POVM $(\Omega_B, \mathscr{B}_B, B)$ weakly fuzzy equivalent to A. Since $B \simeq_w A \simeq_w \tilde{A}$, there exist mappings $f \in \mathbb{M}((\mathbb{R}^\infty, \mathscr{B}(\mathbb{R}^\infty)) \to (\Omega_B, \mathscr{B}_B))$ and $g \in \mathbb{M}((\Omega_B, \mathscr{B}_B) \to (\mathbb{R}^\infty, \mathscr{B}(\mathbb{R}^\infty)))$ such that $\tilde{A}_f = B$ and $B_g = \tilde{A}$. Then, we have $\tilde{A}_{g \circ f} = \tilde{A}$, and from Theorem 3, $g \circ f \colon \mathbb{R}^\infty \to \mathbb{R}^\infty$ is a sufficient statistic for $\tilde{A}$ and we have

$$\frac{dP^{\tilde{A}}_{\rho_i}}{dP^{\tilde{A}}_{\rho_*}}(t) = \frac{dP^{\tilde{A}}_{\rho_i}}{dP^{\tilde{A}}_{\rho_*}}(g \circ f(t)) \quad \tilde{A}\text{-a.e.} \tag{16}$$

for each $i \geq 1$. From Eqs. (15) and (16), we obtain

$$t = \left(\frac{dP^{\tilde{A}}_{\rho_i}}{dP^{\tilde{A}}_{\rho_*}}(t)\right)_{i \geq 1} = \left(\frac{dP^{\tilde{A}}_{\rho_i}}{dP^{\tilde{A}}_{\rho_*}}(g \circ f(t))\right)_{i \geq 1} = g \circ f(t) \quad \tilde{A}\text{-a.s.}$$

Thus, there exists a full measure set $\tilde{\Omega}_{\tilde{A}} \in \mathscr{B}(\mathbb{R}^\infty)$ such that $(g \circ f)|_{\tilde{\Omega}_{\tilde{A}}}$ is an identity map on $\tilde{\Omega}_{\tilde{A}}$. Then, $\tilde{\Omega}_B := f(\tilde{\Omega}_{\tilde{A}}) = g^{-1}(\tilde{\Omega}_{\tilde{A}}) \in \mathscr{B}_B$ is a full measure set and $f|_{\tilde{\Omega}_{\tilde{A}}}$ is a strict isomorphism between $(\tilde{\Omega}_{\tilde{A}}, \tilde{\Omega}_{\tilde{A}} \cap \mathscr{B}(\mathbb{R}^\infty), \tilde{A}|_{\tilde{\Omega}_{\tilde{A}}})$ and $(\tilde{\Omega}_B, \tilde{\Omega}_B \cap \mathscr{B}_B, B|_{\tilde{\Omega}_B})$. Therefore, $\tilde{A}$ and B are almost isomorphic.

2. As shown in 1, the POVM $\tilde{A}$ induced by the LSB statistic for A is $\simeq_w$-minimal sufficient POVM, and therefore $\simeq_r$-minimal sufficient POVM, weakly fuzzy equivalent to A. Since A and $\tilde{A}$ are standard Borel POVMs, they are regularly fuzzy equivalent. The uniqueness up to almost isomorphism can be shown in a similar manner.   □

IV. MINIMAL SUFFICIENCY FOR DISCRETE POVM

In this section, we consider the minimal sufficient condition for discrete POVMs.

A measurable space $(\Omega, \mathscr{B})$ is said to be discrete if Ω is a countable set and $\mathscr{B}$ is the power set $\mathcal{P}(\Omega)$ of Ω. A POVM A is said to be discrete if the outcome space of A is a discrete space. A discrete POVM $(\Omega_A, \mathcal{P}(\Omega_A), A)$ induces a mapping $A \colon \Omega_A \to \mathcal{L}(\mathcal{H})$ defined by $A(x) := A(\{x\}) \geq O$ with a completeness condition

$$\sum_{x \in \Omega_A} A(x) = I. \tag{17}$$

On the other hand, a positive-operator valued mapping $A \colon \Omega_A \to \mathcal{L}(\mathcal{H})$ with completeness condition (17) induces a discrete POVM by

$$A(E) := \sum_{x \in E} A(x)$$

for each $E \in \mathcal{P}(\Omega_A)$. Thus, throughout this section, we identify a positive-operator valued mapping $A \colon \Omega_A \to \mathcal{L}(\mathcal{H})$ satisfying the completeness condition with a discrete POVM. We write $p^A_\rho(x) := \mathrm{tr}[\rho A(x)]$ for each $\rho \in \mathcal{S}(\mathcal{H})$.

A discrete POVM $A \colon \Omega_A \to \mathcal{L}(\mathcal{H})$ is said to be *non-vanishing* if $A(x) \neq 0$ for all $x \in \Omega_A$. Any discrete POVM is almost isomorphic to a non-vanishing POVM.

Since a discrete space is a standard Borel space, the weak relations $\preceq_w$ and $\simeq_w$ coincide with the regular relations $\preceq_r$ and $\simeq_r$, respectively. Thus, for discrete POVMs, fuzzy preorder and equivalence relations are denoted as $\preceq$ and $\simeq$, respectively.

For discrete POVMs $A \colon \Omega_A \to \mathcal{L}(\mathcal{H})$ and $B \colon \Omega_B \to \mathcal{L}(\mathcal{H})$ the relations $\approx$, $\sim$, and $\preceq$ are simplified as follows. $A \approx B$ if and only if there exists a bijection $f \colon \Omega_A \to \Omega_B$ such that $A_f = B$, where $A_f(y) = \sum_{x \in \Omega_A} \delta_{y, f(x)} A(x)$. $A \sim B$ if and only if there exist full-measure subsets $\tilde{\Omega}_A \subset \Omega_A$ and $\tilde{\Omega}_B \subset \Omega_B$ such that the restrictions $A|_{\tilde{\Omega}_A}$ and $B|_{\tilde{\Omega}_B}$ are non-vanishing and strictly isomorphic. $A \preceq B$ if and only if there exists a matrix $\{\kappa(x|y)\}_{(x,y) \in \Omega_A \times \Omega_B}$ such that

$$\kappa(x|y) \geq 0, \quad \sum_{x \in \Omega_A} \kappa(x|y) = 1, \tag{18}$$

$$A(x) = \sum_{y \in \Omega_B} \kappa(x|y) B(y).$$

A matrix $\kappa(\cdot|\cdot)$ satisfying condition (18) is called a Markov matrix.

The following proposition characterizes the sufficiency condition of a statistic for a discrete POVM.

Proposition 5. Let $A \colon \Omega_A \to \mathcal{L}(\mathcal{H})$ be a discrete POVM and let $T \colon \Omega_A \to \Omega_T$ be a mapping to a measurable space $(\Omega_T, \mathcal{B}_T)$. We assume that $\mathcal{B}_T$ contains each single point set $\{t\}(t \in \Omega_T)$. Then, the following conditions are equivalent:

(i) T is sufficient for A;

(ii)   there exist functions $h(\cdot) \colon \Omega_A \ni x \mapsto h(x) \in [0, \infty)$ and $g_{\cdot}(\cdot) \colon \mathcal{S}(\mathcal{H}) \times \Omega_T \ni (\rho, t) \mapsto g_\rho(t) \in [0, \infty)$ such that

$$p_\rho^A(x) = h(x)g_\rho(T(x)), \quad (\rho \in \mathcal{S}(\mathcal{H}), x \in \Omega_A);$$

(iii)   there exist functions $h(\cdot) \colon \Omega_A \ni x \mapsto h(x) \in [0, \infty)$ and $G(\cdot) \colon \Omega_T \ni t \mapsto G(t) \in \mathcal{L}_+(\mathcal{H})$ such that

$$A(x) = h(x)G(T(x)), \quad (x \in \Omega_A),$$

where $\mathcal{L}_+(\mathcal{H}) \coloneqq \{a \in \mathcal{L}(\mathcal{H}) | a \geq O\}$;

(iv) $A \simeq A_T$.

Proof. The equivalence (i)⇔(ii)⇔(iv) is evident from Theorem 3. The implication (iii)⇒(ii) immediately follows by putting $g_\rho(t) \coloneqq \text{tr}[\rho G(t)]$. Let us assume (ii) and show (iii). If $t \in \Omega_T$ satisfies that

$$\exists x \in \Omega_A \text{ such that } T(x) = t \text{ and } h(x) \neq 0, \tag{19}$$

then we have

$$g_\rho(t) = p_\rho^A(x)/h(x). \tag{20}$$

Since the RHS of Eq. (20) is affine and positive with respect to $\rho \in \mathcal{S}(\mathcal{H})$, according to Refs. 2 and 19, there exists $G(t) \in \mathcal{L}_+(\mathcal{H})$ such that $g_\rho(t) = \text{tr}[\rho G(t)]$ for any $\rho \in \mathcal{S}(\mathcal{H})$. For $t \in \Omega_T$ that does not satisfy condition (19), we define $G(t) = O$. Then, we have

$$\text{tr}[\rho A(x)] = p_\rho^A(x) = \text{tr}[\rho h(x)G(T(x))]$$

for each $\rho \in \mathcal{S}(\mathcal{H})$ and $x \in \Omega_A$, which implies condition (iii). □

Corresponding to Theorem 5, we have the following theorem as to the existence and uniqueness of a minimal sufficient POVM.

Theorem 6. *Let $A \colon \Omega_A \to \mathcal{L}(\mathcal{H})$ be a discrete POVM. Then, there exists a discrete non-vanishing $\simeq_w$-minimal sufficient POVM $\bar{A}$ fuzzy equivalent to A. Furthermore, such $\bar{A}$ is unique up to strict isomorphism.*

Proof. Since A is almost isomorphic to a non-vanishing POVM, without loss of generality, we can assume that A is non-vanishing. We define an equivalence relation on Ω_A by

$$x \sim_A x' \colon \Leftrightarrow \exists c > 0, A(x) = cA(x')$$

and define a mapping $S \colon \Omega_A \to \Omega_A / \sim_A =: \Omega_S$ by $S(x) \coloneqq [x]$, where $[x]$ is the equivalence class to which x belongs. Then, from the definition of $\sim_A$, we can write

$$A(x) = h(x)G(S(x))$$

for each $x \in \Omega_A$, where $h(x) > 0$ and $O \neq G(s) \in \mathcal{L}_+(\mathcal{H})(s \in \Omega_S)$, which implies S is sufficient for A. Therefore, if we define $\bar{A} \coloneqq A_S$, $\bar{A} \simeq A$ and $\bar{A}$ is a non-vanishing discrete POVM.

From Theorem 5 and its proof, the LSB statistic $T \colon \Omega_S \to \mathbb{R}^\infty$ defined by Eq. (6) induces a $\simeq_w$-minimal sufficient POVM $(\mathbb{R}^\infty, \mathcal{B}(\mathbb{R}^\infty), \bar{A}_T)$ fuzzy equivalent to $\bar{A}$. Since T is sufficient for $\bar{A}$, from Proposition 5, we can write

$$\bar{A}(s) = h'(s)G'(T(s)),$$

where $h'(s) > 0$ and $O \neq G'(t) \in \mathcal{L}_+(\mathcal{H})(t \in \mathbb{R}^\infty)$. From the construction of S, this implies that T is injective, and T is a strict isomorphism between $(\Omega_S, \mathcal{P}(\Omega_S), \bar{A})$ and $(T(\Omega_S), T(\Omega_S) \cap \mathcal{B}(\mathbb{R}^\infty), (\bar{A}_S)|_{T(\Omega_S)})$. Note that $T(\Omega_S) \in \mathcal{B}(\mathbb{R}^\infty)$ since $T(\Omega_S)$ is a countable set. Thus, from Proposition 4, $\bar{A}$ is a $\simeq_w$-minimal sufficient POVM.

To show the uniqueness, let $B\colon \Omega_B \to \mathcal{L}(\mathcal{H})$ be an arbitrary $\simeq_w$-minimal sufficient and non-vanishing discrete POVM equivalent to A. Since $\bar{A} \simeq B$, there exists a mapping $f\colon \Omega_S \to \Omega_B$ such that $A_f = B$. If f is not surjective, there exists an element $y \in \Omega_B \setminus f(\Omega_S)$ and we have $B(y) = A_f(y) = \sum_{x\colon f(x)=y} A(x) = O$, which contradicts the non-vanishing property of B. Thus, f is surjective. Since f is sufficient for $\bar{A}$, by a similar discussion for T, we can show that f is injective. Therefore, f is a strict isomorphism between $\bar{A}$ and B, which completes the proof. $\qquad\square$

A discrete POVM $A\colon \Omega_A \to \mathcal{L}(\mathcal{H})$ is said to be pairwise linearly independent[5] if any pair $\{A(x), A(x')\}(x \neq x')$ is linearly independent. A pairwise linearly independent POVM is non-vanishing by definition. The following theorem states that the minimal sufficiency and the pairwise linearly independence are almost equivalent.

Theorem 7. *Let $A\colon \Omega_A \to \mathcal{L}(\mathcal{H})$ be a discrete POVM. Then, the following conditions are equivalent:*

- *(i) A is pairwise linearly independent;*
- *(ii) A is non-vanishing and $\simeq_w$-minimal sufficient;*
- *(iii) A is non-vanishing and $\simeq_r$-minimal sufficient.*

Proof. We first show (i)$\Rightarrow$(ii). Assume that A is a pairwise linearly independent POVM. Then if we consider the mapping S in the proof of Theorem 6, S is an injection and the POVM $\bar{A}$ induced by S is strictly isomorphic to A. Since $\bar{A}$ is $\simeq_w$-minimal sufficient, A is also $\simeq_w$-minimal sufficient. Thus, we have shown (i)$\Rightarrow$(ii).

(ii)$\Rightarrow$(iii) is obvious.

If A is non-vanishing and $\simeq_r$-minimal sufficient, then $\bar{A}$ induced by the statistic S is strictly isomorphic to A by the uniqueness of the minimal sufficient POVM. Since $\bar{A}$ is pairwise linearly independent, A is also pairwise linearly independent. $\qquad\square$

In Ref. 5, it is shown that for each discrete POVM A, there exists a pairwise linearly independent POVM fuzzy equivalent to A and such a POVM is unique up to strict isomorphism. This assertion is a direct corollary of our Theorems 6 and 7.

V. INFORMATION CONSERVATION CONDITION

In this brief section, we consider information conservation conditions proposed by the author.[9,10]

Let $(\Omega_1, \mathscr{B}_1)$ be a measurable space. A completely positive (CP) instrument[20] $\mathcal{I}^1(\cdot)$ (in the Heisenberg picture) with the outcome space $(\Omega_1, \mathscr{B}_1)$ is a mapping

$$\mathcal{I}^1(\cdot)\colon \mathscr{B}_1 \times \mathcal{L}(\mathcal{H}) \ni (E_1, a) \mapsto \mathcal{I}^1_{E_1}(a) \in \mathcal{L}(\mathcal{H})$$

such that

- (i)   for each countable disjoint $\{E_j\} \subset \mathscr{B}_1$ and each $\rho \in \mathcal{S}(\mathcal{H})$ and $a \in \mathcal{L}(\mathcal{H})$, $\mathrm{tr}[\rho \mathcal{I}^1_{\cup_j E_j}(a)] = \sum_j \mathrm{tr}[\rho \mathcal{I}^1_{E_j}(a)]$;
- (ii) $\mathcal{I}_{\Omega_1}(I) = I$;
- (iii) $\mathcal{I}^1_E(\cdot)$ is a normal CP linear map for every $E \in \mathscr{B}_1$.

A CP instrument simultaneously describes the probability distribution of the outcome of a quantum measurement process and the state change due to the measurement.

Let $\mathcal{I}^1(\cdot)$ be a CP instrument with a standard Borel outcome space $(\Omega_1, \mathscr{B}_1)$ and let $(\Omega_2, \mathscr{B}_2, A_2)$ be a standard Borel POVM. A composition[2,10] $\mathcal{I}^1 * A_2$ is a unique POVM with the product outcome space $(\Omega_1 \times \Omega_2, \mathscr{B}_1 \times \mathscr{B}_2)$ such that

$$(\mathcal{I}^1 * A_2)(E_1 \times E_2) = \mathcal{I}^1_{E_1}(A_2(E_2))$$

for each $E_1 \in \mathscr{B}_1$ and $E_2 \in \mathscr{B}_2$. The composition corresponds to the joint successive measurement process of $\mathcal{I}^1(\cdot)$ followed by A_2.

For a given CP instrument $\mathcal{I}^1(\cdot)$ with a standard Borel outcome space and a given standard Borel POVM A_2, we consider the following two conditions:

1. There exists a sufficient statistic $\tilde{x}\colon \Omega_1 \times \Omega_2 \to \Omega_2$ such that $(\mathcal{I}^1 * A_2)_{\tilde{x}} = A_2$;
2. $\mathcal{I}^1 * A_2 \simeq_r A_2$.

In Ref. 9, the author derived condition 1 as a sufficient condition for a so called relative-entropy conservation law. In Ref. 10, the author reformulated this condition in the form of 2 and called it an information conservation condition, while condition 1 is a sufficient condition but not a necessary one for condition 2. Noting that condition 2 is invariant under replacing A_2 with another regularly fuzzy equivalent standard Borel POVM,[10] this discrepancy can be resolved by taking A_2 an equivalent $\simeq_w$-minimal sufficient POVM, which is always possible due to Theorem 5, and in this sense, the two conditions are essentially equivalent.

ACKNOWLEDGMENTS

APPENDIX: INEQUIVALENCE OF $\simeq_r$ AND $\simeq_w$

In this appendix, we construct a pair of POVMs that are weakly fuzzy equivalent but not regularly fuzzy equivalent.

Let $(\Omega_A, \mathcal{B}_A, A)$ be a POVM. The completion of $(\Omega_A, \mathcal{B}_A, A)$ is a POVM $(\Omega_A, \bar{\mathcal{B}}_A, \bar{A})$ defined by

$$\mathcal{N}_A := \{N \subset \Omega_A \mid \exists N' \in \mathcal{B}_A \text{ s.t. } N \subset N' \text{ and } A(N') = O\},$$
$$\bar{\mathcal{B}}_A := \{E \subset \Omega_A \mid \exists F \in \mathcal{B}_A \text{ s.t. } E \triangle F \in \mathcal{N}_A\},$$
$$\bar{A}(E) := A(F), \quad (E \in \bar{\mathcal{B}}_A, F \in \mathcal{B}_A, E \triangle F \in \mathcal{N}_A).$$

Here, $E \triangle F := (E \setminus F) \cup (F \setminus E)$ is the symmetric difference of sets. An element of $\mathcal{N}_A$ is called an A-null set.

Lemma 3. Let $(\Omega_A, \mathcal{B}_A, A)$ be a POVM and let $(\Omega_A, \bar{\mathcal{B}}_A, \bar{A})$ be the completion of $(\Omega_A, \mathcal{B}_A, A)$. Then $A \simeq_w \bar{A}$.

Proof. Since $A(E) = \bar{A}(E) = \int_{\Omega_A} \chi_E(x) d\bar{A}(x)$ for each $E \in \mathcal{B}_A$, $A \preceq_r \bar{A}$ holds. For each $E \in \bar{\mathcal{B}}_A$, we take $F \in \mathcal{B}_A$ such that $E \triangle F \in \mathcal{N}_A$ and define $\kappa(E|x) := \chi_F(x)$. Then, $\kappa(\cdot|\cdot)$ is an A-weak Markov kernel such that $\bar{A}(E) = \int_{\Omega_A} \kappa(E|x) dA(x)$ for each $E \in \bar{\mathcal{B}}_A$. Thus, we have $\bar{A} \preceq_w A$, and the assertion holds.   □

Let μ be the usual Lebesgue measure on a unit interval $I := [0,1]$, i.e., μ is the unique measure defined on the σ-algebra $\mathcal{B}(I)$ of I generated by open subsets of I such that $\mu([a,b]) = b - a$ for each $0 \le a < b \le 1$. As the system Hilbert space $\mathcal{H}$, we consider the set of square-integrable $\mathcal{B}(I)$-measurable functions $L^2(I, \mathcal{B}(I), \mu)$ in which μ-a.e. equal functions are identified. We define a projection-valued measure (PVM) $(I, \mathcal{B}(I), A)$ by

$$(A(E)f)(x) := \chi_E(x)f(x)$$

for each $E \in \mathcal{B}(I)$ and $f \in \mathcal{H}$. We denote the completion of $(I, \mathcal{B}(I), A)$ by $(I, \bar{\mathcal{B}}(I), \bar{A})$. Since the class of A-null sets and that of μ-null sets coincide, $\bar{\mathcal{B}}(I)$ is the class of Lebesgue measurable subsets of I. From Lemma 3, we have $A \simeq_w \bar{A}$.

Now we prove that $\bar{A} \not\preceq_r A$, from which we immediately obtain the desired relation $A \not\simeq_r \bar{A}$. Suppose that $\bar{A} \preceq_r A$ holds. Then, there exists a regular Markov kernel $\kappa(\cdot|\cdot)\colon \bar{\mathcal{B}}(I) \times I \to [0,1]$ such that

$$\bar{A}(E) = \int_I \kappa(E|x) dA(x)$$

for each $E \in \bar{\mathcal{B}}(I)$. Thus, we have

$$\int_I \chi_E(x) dA(x) = A(E) = \bar{A}(E) = \int_I \kappa(E|x) dA(x) \tag{A1}$$

for each $E \in \mathscr{B}(I)$. From Remark 5 of Ref. 6, Eq. (A1) implies that $\chi_E(x) = \kappa(E|x)$ for μ-a.e. $x \in I$. Therefore, there exists a μ-null set $N \in \mathscr{B}(I)$ such that

$$\kappa([0,r]|x) = \chi_{[0,r]}(x) \quad (\forall r \in I \cap \mathbb{Q}, \forall x \in I \setminus N), \tag{A2}$$

where $\mathbb{Q}$ is the set of rational numbers. Noting that $\kappa(\cdot|x)$ is a probability measure for each $x \in I$, Eq. (A2) indicates that $\kappa(E|x) = \chi_E(x)$ for each $x \in I \setminus N$ and $E \in \mathscr{B}(I)$. Thus, we have $\kappa(I \setminus \{x\}|x) = 0$ for each $x \in I \setminus N$ and this implies that $\kappa(E|x) = \chi_E(x)$ for each $x \in I \setminus N$ and $E \in \bar{\mathscr{B}}(I)$. If there exists a set E such that

$$E \subset I \setminus N, \quad E \in \bar{\mathscr{B}}(I) \setminus \mathscr{B}(I), \tag{A3}$$

then we have $\chi_E(x) = \kappa(E|x)\chi_{I \setminus N}(x)$, which contradicts the $\mathscr{B}(I)$-nonmeasurability of $\chi_E(x)$.

Now we show the existence of E satisfying (A3). Since $(I \setminus N, (I \setminus N) \cap \mathscr{B}(I))$ is a standard Borel space and the restriction of μ to $I \setminus N$ is a continuous measure, from Theorem 17.41 of Ref. 21, there exists a $(I \setminus N) \cap \mathscr{B}(I)/\mathscr{B}(I)$-bimeasurable bijection $f : I \setminus N \to I$ such that $\mu(f(E)) = \mu(E)$ for every $E \in (I \setminus N) \cap \mathscr{B}(I)$. Since there exists a set $\tilde{E} \in \bar{\mathscr{B}}(I) \setminus \mathscr{B}(I)$, $E := f^{-1}(\tilde{E})$ satisfies condition (A3), which completes the proof of the assertion.

[1] E. B. Davies and J. Lewis, "An operational approach to quantum probability," Commun. Math. Phys. **17**, 239–260 (1970).

[2] E. B. Davies, *Quantum Theory of Open Systems* (Academic Press, London, 1976).

[3] E. L. Lehmann and H. Scheffé, "Completeness, similar regions, and unbiased estimation: Part I," Sankhyā **10**, 305–340 (1950).

[4] R. R. Bahadur, "Sufficiency and statistical decision functions," Ann. Math. Stat. **25**, 423–462 (1954).

[5] H. Martens and W. de Muynck, "Nonideal quantum measurements," Found. Phys. **20**, 255–281 (1990).

[6] S. Dorofeev and J. de Graaf, "Some maximality results for effect-valued measures," Indagationes Math. **8**, 349–369 (1997).

[7] T. Heinonen, "Optimal measurements in quantum mechanics," Phys. Lett. A **346**, 77–86 (2005).

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

[9] Y. Kuramochi and M. Ueda, "Classicality condition on a system observable in a quantum measurement and a relative-entropy conservation law," Phys. Rev. A **91**, 032110 (2015).

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

[11] P. R. Halmos and L. J. Savage, "Application of the Radon-Nikodym theorem to the theory of sufficient statistics," Ann. Math. Stat. **20**, 225–241 (1949).

[12] I. Csiszár, "Information-type measures of difference of probability distributions and indirect observations," Stud. Sci. Math. Hung. **2**, 299–318 (1967).

[13] F. Liese and I. Vajda, "On divergences and informations in statistics and information theory," IEEE Trans. Inf. Theory **52**, 4394–4412 (2006).

[14] S. Kullback and R. A. Leibler, "On information and sufficiency," Ann. Math. Stat. **22**, 79–86 (1951).

[15] A. Jenčová and S. Pulmannová, "Characterizations of commutative POV measures," Found. Phys. **39**, 613–624 (2009).

[16] S. M. Srivastava, *A Course on Borel Sets* (Springer, 1998).

[17] K. Itô, *An Introduction to Probability Theory* (Cambridge University Press, 1984).

[18] V. Bogachev, *Measure Theory* (Springer, 2007).

[19] T. Heinosaari and M. Ziman, *The Mathematical Language of Quantum Theory: From Uncertainty to Entanglement* (Cambridge University Press, 2011).

[20] M. Ozawa, "Quantum measuring processes of continuous observables," J. Math. Phys. **25**, 79–87 (1984).

[21] A. Kechris, *Classical Descriptive Set Theory* (Springer, 1995).