Quantum postselective measurements: Sufficient condition for overcoming the Holevo bound and the role of max-relative entropy

N. R. Kenbaev [1,2] and D. A. Kronberg [3,*]

[1] *Terra Quantum AG, St. Gallerstrasse 16A, CH-9400 Rorschach, Switzerland*
[2] *Moscow Institute of Physics and Technology, Institutskii Per. 9, Dolgoprudny 141700, Russia*
[3] *Steklov Mathematical Institute of Russian Academy of Sciences, Gubkina St. 8, Moscow 119991, Russia*

(Received 22 June 2021; revised 7 October 2021; accepted 22 December 2021; published 20 January 2022)

The operation of unambiguous state discrimination (USD) plays an important role in quantum information theory. Here, we study the generalization for this operation which can be applied where USD is not available, including the case of completely coinciding supports. We demonstrate the effect of overcoming the Holevo bound after postselective measurements, which works for a class of ensembles, including two arbitrary noncommuting quantum states. We also discuss the role of max-relative entropy in postselective measurements and give an alternative operational interpretation for this function.

DOI: 10.1103/PhysRevA.105.012609

I. INTRODUCTION

The problem of discrimination between quantum states [1–6] plays an important role in quantum communication and quantum cryptography. There are two main optimization tasks when finding the best measurement: minimizing average error probability and maximizing mutual information, which quantifies how much information can be transmitted between the sender and the receiver.

When mutual information is estimated, the key result is the Holevo bound [7], which states that when quantum ensemble $\{p_i, \hat{\rho}_i\}_{i=1}^N$ of states $\{\hat{\rho}_i\}$ with *a priori* probabilities $\{p_i\}$ is measured, mutual information between input and output is upper bounded by the Holevo quantity:

$$\chi(\{p_i, \hat{\rho}_i\}) = H(\hat{\rho}) - \sum_{i=1}^{N} p_i H(\hat{\rho}_i),$$

$$\hat{\rho} = \sum_{i=1}^{N} p_i \hat{\rho}_i, \quad H(\hat{\rho}) = -\mathrm{Tr}\hat{\rho}\log_2\hat{\rho}, \tag{1}$$

Thus, increasing the number of quantum states cannot lead to infinite mutual information.

We must note that the bound (1) holds only if the users cannot perform postselection, i.e., they must use all the measurement results. But when they can discard some positions, the Holevo bound does not generally hold. The important example for this phenomena is unambiguous state discrimination (USD) [8–14], whether it provides full information (i.e., mutual information is perfect) or yields an inconclusive result. But, USD is possible not for all ensembles of quantum states; see [15].

The motivation for this work is to design a generalization of unambiguous state discrimination which can increase mutual information when USD is not available. We demonstrate the effect of overcoming the Holevo bound for postselective measurements, the phenomena which also works in the case of completely coinciding supports. In Sec. II, we provide a simple example of two qubits. Then, in Sec. III we provide a sufficient condition for postselective mutual information to overcome the Holevo bound; we also show this effect for two arbitrary noncommuting quantum states and give the upper bound for postselective mutual information. This operation has a similarity to maximum confidence quantum measurement [16], and we compare these methods in Sec. IV. We also discuss the role of max-relative entropy in postselective measurements and give an alternative operational interpretation for this function.

II. TWO-QUBIT EXAMPLE

Let us first describe the main idea with a simple example [17]. Consider two arbitrary qubit states $\hat{\rho}_1$ and $\hat{\rho}_2$. They can be expressed as (see Fig. 1)

$$\begin{aligned}\hat{\rho}_1 &= (1-\alpha_1)|\phi_1\rangle\langle\phi_1| + \alpha_1|\phi_2\rangle\langle\phi_2|,\\ \hat{\rho}_2 &= \alpha_2|\phi_1\rangle\langle\phi_1| + (1-\alpha_2)|\phi_2\rangle\langle\phi_2|,\end{aligned} \tag{2}$$

for some positive $\{\alpha_1, \alpha_2\}$. If $[\hat{\rho}_1, \hat{\rho}_2] \neq 0$, the states $|\phi_1\rangle$ and $|\phi_2\rangle$ are not orthogonal. Let us now consider unambiguous state discrimination for the states $\{|\phi_1\rangle, |\phi_2\rangle\}$, where success probability is the same for both states. This operation in the case of success maps them to the orthogonal states $\{|e_1\rangle, |e_2\rangle\}$ which correspond to classical outcomes, thus the states $\{\hat{\rho}_1, \hat{\rho}_2\}$ are mapped onto

$$\begin{aligned}\hat{\rho}_1' &= (1-\alpha_1)|e_1\rangle\langle e_1| + \alpha_1|e_2\rangle\langle e_2|,\\ \hat{\rho}_2' &= \alpha_2|e_1\rangle\langle e_1| + (1-\alpha_2)|e_2\rangle\langle e_2|.\end{aligned} \tag{3}$$

Now observe that the quantum channel,

$$\Phi[\hat{\rho}] = |\phi_1\rangle\langle e_1|\hat{\rho}|e_1\rangle\langle\phi_1| + |\phi_2\rangle\langle e_2|\hat{\rho}|e_2\rangle\langle\phi_2|, \tag{4}$$

*Corresponding author: dmitry.kronberg@gmail.com

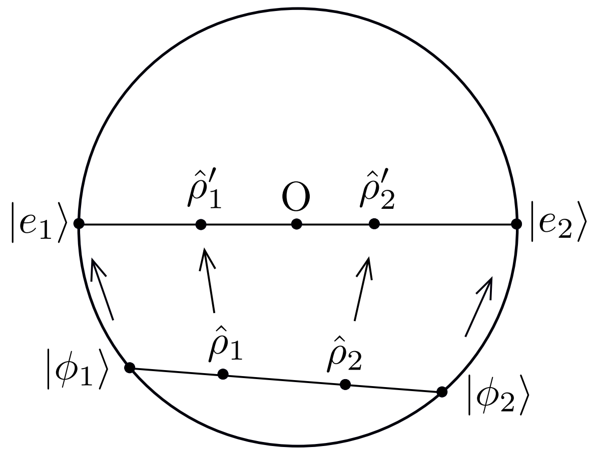

FIG. 1. The expression for two mixed qubits $\hat{\rho}_1$ and $\hat{\rho}_2$ through a convex combination of extreme pure states $|\phi_1\rangle$ and $|\phi_2\rangle$, shown in a Bloch sphere. After unambiguous discrimination between $|\phi_1\rangle$ and $|\phi_2\rangle$, the states $\hat{\rho}_1$ and $\hat{\rho}_2$ pass into the states $\hat{\rho}_1'$ and $\hat{\rho}_2'$.

maps $\{\hat{\rho}_1', \hat{\rho}_2'\}$ onto $\{\hat{\rho}_1, \hat{\rho}_2\}$, thus we can conclude that the Holevo quantity of the states $\{\hat{\rho}_1', \hat{\rho}_2'\}$ is larger than the Holevo quantity for original states $\{\hat{\rho}_1, \hat{\rho}_2\}$, due to its monotonicity after any channel, including (4). Thus we have shown that classical mutual information provided by USD outcomes $\{|e_1\rangle, |e_2\rangle\}$, which is given by the Holevo quantity of the states $\{\hat{\rho}_1', \hat{\rho}_2'\}$, overcomes the Holevo bound of the original states. This effect takes place for two arbitrary noncommuting qubits. If $\hat{\rho}_1$ and $\hat{\rho}_2$ commute, this operation does nothing. Note that for mixed qubits USD is not available, but overcoming the Holevo bound takes place. In the next section, we consider a more general case.

Figure 1 shows on the Bloch sphere how this operation works. In the sequel, we will also discuss geometrical interpretation of the proposed measurement.

III. POSTSELECTIVE INFORMATION AND OVERCOMING THE HOLEVO BOUND

Now we are going to generalize the qubit example from the previous section. First, let us define the mutual information after postselection.

Let $\mathcal{B}(\mathcal{H})$ be the algebra of linear operators acting on a finite-dimensional Hilbert space $\mathcal{H}$ of dimension $d = \dim \mathcal{H}$. Let $\{p_i, \hat{\rho}_i\}_{i=1}^N$ be an ensemble of N states $\{\hat{\rho}_i\}$ in $\mathcal{B}(\mathcal{H})$ with *a priori* probabilities $\{p_i\}$.

Next, let $\Pi = \{\hat{\Pi}_k\}_{k=0}^K$ be POVM, where $\hat{\Pi}_0$ corresponds to inconclusive result. When measuring $\hat{\rho}_i$, the probability of outcome k is $p(k|i) = \mathrm{Tr}\hat{\rho}_i\hat{\Pi}_k$. We also consider a restriction,

$$p(0|i) = 1 - \sum_{k=1}^{K} \mathrm{Tr}\hat{\rho}_i\hat{\Pi}_k = p_? < 1 \quad \forall i, \tag{5}$$

i.e., inconclusive probability is the same for every state $\hat{\rho}_i$. It means that postselective measurement should not change *a priori* state probabilities $\{p_i\}$ in the case of success or fail. Actually, one can omit this restriction, but we think it is reasonable, as we describe the *quantum* effect of postselection instead of just the classical effect of changing signal probabilities, which of course can also increase mutual information.

Next, let us define conditional probabilities in the case of a conclusive result,

$$p_{ps}(k|i) = \frac{\mathrm{Tr}\hat{\rho}_i\hat{\Pi}_k}{\sum_{k'=1}^{K}\mathrm{Tr}\hat{\rho}_i\hat{\Pi}_{k'}} = \frac{p(k|i)}{1-p_?},$$

and mutual information when a conclusive result is obtained:

$$I_{ps}^{\Pi}(\{p_i, \hat{\rho}_i\}) = \sum_{i,k} p_i p_{ps}(k|i) \log_2 \frac{p_{ps}(k|i)}{\sum_j p_{ps}(k|j)p_j}.$$

When the best postselective measurement is performed, we have

$$I_{ps}(\{p_i, \hat{\rho}_i\}) = \max_{\Pi} I_{ps}^{\Pi}(\{p_i, \hat{\rho}_i\}). \tag{6}$$

There are two trivial bounds for $I_{ps}(\{p_i, \hat{\rho}_i\})$:

$$\max_M I_1(M, \{p_i, \hat{\rho}_i\}) \leqslant I_{ps}(\{p_i, \hat{\rho}_i\}) \leqslant H(\{p_i\}), \tag{7}$$

where $I_1(M, \{p_i, \hat{\rho}_i\})$ is one-shot mutual information, i.e., mutual information when non-postselective observable M is applied (see, e.g., [18]), and $H(\{p_i\})$ is the Shannon entropy for signal probability distribution. Our goal is to improve these bounds.

We are now ready to formulate the main result for overcoming the Holevo bound with postselective measurements.

Theorem 1. Suppose that $\hat{\rho}_i$ can be expressed as a convex combination of the given set $\{\hat{\sigma}_k\}_{k=1}^K$ of extreme states,

$$\hat{\rho}_i = \sum_{k=1}^{K} \alpha_{ik}\hat{\sigma}_k, \quad i = 1, \dots, N, \tag{8}$$

and $\{\hat{\sigma}_k\}_{k=1}^K$ can be unambiguously discriminated. Then

$$I_{ps}(\{p_i, \hat{\rho}_i\}) \geqslant \chi(\{p_i, \hat{\rho}_i\}). \tag{9}$$

Proof. Since states $\hat{\sigma}_k$ can be unambiguously discriminated, there exists a state $|\psi_k\rangle$ such that $\langle\psi_k|\hat{\sigma}_k|\psi_k\rangle > 0$ but $\langle\psi_k|\hat{\sigma}_j|\psi_k\rangle = 0$ for any $j \neq k$. Then, the unambiguous state discrimination for extreme states is defined by POVM elements,

$$\hat{\Pi}_k = c_k|\psi_k\rangle\langle\psi_k|, \quad k = 1, \dots, K,$$
$$\hat{\Pi}_0 = \hat{I} - \sum_{k=1}^{K}\hat{\Pi}_k, \tag{10}$$

where $\hat{\Pi}_0$ corresponds to an inconclusive outcome. One can also consider multidimensional projectors here but for our demonstration it is redundant.

To ensure restriction (5) we choose the parameters c_k in such a way that the probability of inconclusive outcome is the same for each extreme state,

$$c_k = \frac{1 - p_?}{\langle\psi_k|\hat{\sigma}_k|\psi_k\rangle}, \quad k = 1, \dots, K. \tag{11}$$

The minimum value of $p_?$ is obtained when the minimum eigenvalue of $\hat{\Pi}_0$ becomes zero.

From such a choice of parameters (11) we directly obtain that conditional probabilities are

$$p_{ps}(k|i) = \alpha_{ik}, \tag{12}$$

and the initial ensemble transforms into

$$\hat{\rho}_i' = \sum_{k=1}^{K} \alpha_{ik} |e_k\rangle\langle e_k|, \tag{13}$$

where $\{|e_k\rangle\}_{k=1}^{K}$ form an orthonormal basis in output space corresponding to classical measurement outcomes.

The inverse mapping to our measurement with postselection can be written as

$$\Phi[\hat{\rho}] = \sum_{k,j} \sigma_{kj} |\phi_{kj}\rangle\langle e_k|\hat{\rho}|e_k\rangle\langle\phi_{kj}|, \tag{14}$$

where $\hat{\sigma}_k = \sum_j \sigma_{kj}|\phi_{kj}\rangle\langle\phi_{kj}|$ is spectral decomposition for $\hat{\sigma}_k$.

Due to the monotonicity of the Holevo quantity under the action of the linear completely positive trace-preserving (CPTP) map, and due to the fact that the states $\{\hat{\rho}_i'\}_{i=1}^{N}$ are diagonal in the basis $\{|e_k\rangle\}_{k=1}^{K}$ we obtain (9). ■

For example, if the states $\{\hat{\sigma}_k\}$ are pure linearly independent states, the theorem condition holds.

This theorem shows that for some quantum ensembles individual postselective measurements can provide higher information than collective measurements without postselection. Recall that when the states do not commute, the Holevo quantity can be achieved only with collective measurements [18,19].

Also, let us once more note that without the restriction (5) the inequality (9) remains correct, and can be applied to a broader class of ensembles.

Note that (9) may also hold in the classical case, which corresponds to commuting states. Consider two equiprobable states in a three-dimensional Hilbert space:

$$\begin{aligned}\hat{\rho}_1^{cl} &= (1-p)|e_1\rangle\langle e_1| + p|e_3\rangle\langle e_3|,\\ \hat{\rho}_2^{cl} &= (1-p)|e_2\rangle\langle e_2| + p|e_3\rangle\langle e_3|,\end{aligned} \tag{15}$$

where $0 < p < 1$. Here, the Holevo quantity is $1-p$. After measuring in the basis $\{|e_1\rangle, |e_2\rangle, |e_3\rangle\}$ and discarding the last outcome we have $I_{ps} = 1$. This simple example shows that the effect of overcoming the Holevo bound is not only quantum.

Nevertheless, overcoming the Holevo bound is possible not for any ensemble. Consider three equiprobable symmetric qubit states:

$$\begin{aligned}|\xi_1\rangle &= |0\rangle,\\ |\xi_2\rangle &= -\frac{1}{2}|0\rangle + \frac{\sqrt{3}}{2}|1\rangle,\\ |\xi_3\rangle &= -\frac{1}{2}|0\rangle - \frac{\sqrt{3}}{2}|1\rangle.\end{aligned} \tag{16}$$

Let us now show that for this ensemble one cannot achieve the Holevo bound with individual measurements, even when postselection is allowed. $I_{ps}^{\Pi}(\{p_i, \hat{\rho}_i\})$ can be rewritten as

$$\begin{aligned}I_{ps}^{\Pi}(\{p_i, \hat{\rho}_i\}) &= H(X) - \sum_{k=1}^{K} p(k) H(X|Y=k)\\ &= H(\{p_i\}) + \sum_{k=1}^{K} p(k) \sum_i p(i|k)\log_2 p(i|k),\end{aligned} \tag{17}$$

where the summation over k is performed over conclusive outcomes only. Now consider the arbitrary rank-one POVM element $\hat{\Pi}_k = |\eta\rangle\langle\eta|$, where

$$|\eta\rangle = \begin{pmatrix}\cos(\theta)\\ \sin(\theta)\end{pmatrix}. \tag{18}$$

Using Bayes' rule, we find that

$$\begin{aligned}H(X|Y=k) = \log_2(3) - \frac{1}{3}\sum_{j=0}^{2}\left(1+\cos\left(2\theta - \frac{2j\pi}{3}\right)\right)\\ \times \log_2\left(1+\cos\left(2\theta - \frac{2j\pi}{3}\right)\right).\end{aligned} \tag{19}$$

According to Lemma 3 in [20], $H(X|Y=k)$ has a global minimum at $\theta = \frac{\pi}{2}$, thus $H(X|Y=k) \geqslant 1$, which implies

$$I_{ps}\left(\left\{\tfrac{1}{3}, |\xi_i\rangle\langle\xi_i|\right\}\right) \leqslant \log_2(3) - 1, \tag{20}$$

whereas the Holevo quantity for this ensemble of three pure states equals 1. This example shows that collective measurements without postselection, which can be used to achieve the Holevo bound when appropriate coding is performed, may be more efficient than individual postselective measurements.

Below we will provide the upper bound for $I_{ps}(\{p_i, \hat{\rho}_i\})$. Notice, that in terms of the quantum relative entropy introduced by Umegaki [21],

$$S(\hat{\rho}_1||\hat{\rho}_2) = \mathrm{Tr}(\hat{\rho}_1\log_2\hat{\rho}_1 - \hat{\rho}_1\log_2\hat{\rho}_2), \tag{21}$$

the expression for the Holevo quantity (1) can be rewritten as follows:

$$\chi(\{p_i, \hat{\rho}_i\}) = \sum_{i=1}^{N} p_i S(\hat{\rho}_i||\hat{\rho}). \tag{22}$$

Max-relative entropy for two operators $\hat{\rho}_1$ and $\hat{\rho}_2$ is defined as [22]

$$D_{\max}(\hat{\rho}_1||\hat{\rho}_2) = \log_2 \min\{\delta : \hat{\rho}_1 \leqslant \delta\hat{\rho}_2\}, \tag{23}$$

which is equivalent to

$$D_{\max}(\hat{\rho}_1||\hat{\rho}_2) = -\log_2 \max\{\lambda : \hat{\rho}_2 - \lambda\hat{\rho}_1 \geqslant 0\}. \tag{24}$$

Theorem 2. Suppose that the restriction (5) holds. Then, mutual information with postselection satisfies the following bound:

$$I_{ps}(\{p_i, \hat{\rho}_i\}) \leqslant \sum_{i=1}^{N} p_i D_{\max}(\hat{\rho}_i||\hat{\rho}). \tag{25}$$

Proof. Let $\hat{\rho}_i'$ be transformed states after the measurement with postselection with associated probability distribution

$\{p_i'\}$. The restriction (5) yields $p_i' = p_i$. Let us define $\hat{\rho}' = \sum_{i=1}^{N} p_i' \hat{\rho}_i'$.

From the monotonicity of max-relative entropy after any completely positive (not necessarily trace-preserving) map (see [23]),

$$D_{\max}(\hat{\rho}_i'||\hat{\rho}') \leqslant D_{\max}(\hat{\rho}_i||\hat{\rho}), \tag{26}$$

we obtain

$$I_{ps}(\{p_i, \hat{\rho}_i\}) = \sum_{i=1}^{N} p_i S(\hat{\rho}_i'||\hat{\rho}') \leqslant \sum_{i=1}^{N} p_i D_{\max}(\hat{\rho}_i'||\hat{\rho}') \leqslant \sum_{i=1}^{N} p_i D_{\max}(\hat{\rho}_i||\hat{\rho}). \tag{27}$$

■

If states of the ensemble $\{p_i, \hat{\rho}_i\}$ allow for unambiguous discrimination, the upper bound in (25) is attained.

Now let us consider the case of two states $\{\hat{\rho}_1, \hat{\rho}_2\}$, not necessarily qubits. The decomposition (8) is not unique, but let us consider the following decomposition similar to (2):

$$\begin{aligned} \hat{\rho}_1 &= (1-\alpha_1)\hat{\sigma}_1 + \alpha_1\hat{\sigma}_2, \\ \hat{\rho}_2 &= \alpha_2\hat{\sigma}_1 + (1-\alpha_2)\hat{\sigma}_2. \end{aligned} \tag{28}$$

Here, $\hat{\sigma}_1$ and $\hat{\sigma}_2$ are not necessarily pure, but for the possibility of USD between $\hat{\sigma}_1$ and $\hat{\sigma}_2$, they must have lower rank than $\hat{\rho} = p_1\hat{\rho}_1 + p_2\hat{\rho}_2$ (see [15]).

Thus, $\{\hat{\sigma}_1, \hat{\sigma}_2\}$ have the form,

$$\hat{\sigma}_i = \frac{\hat{\rho}_i - \lambda_i\hat{\rho}}{1-\lambda_i}, \quad i \in \{1, 2\}, \tag{29}$$

where $\lambda_i \in \mathbb{R}$ is the minimal real number, for which $\hat{\rho}_i - \lambda_i\hat{\rho}$ has lower rank than $\hat{\rho}$. Some $\hat{\sigma}_i$ may also coincide with $\hat{\rho}_i$, if the rank of $\hat{\rho}_1$ differs from the rank of $\hat{\rho}_2$.

This decomposition demonstrates that overcoming the Holevo bound with postselective measurement takes place for any two noncommuting states, not only qubits.

Note that here we do not claim the optimality of decompositions (8) or (28) for maximization of postselective mutual information (6).

Decomposition (28) is closely connected with the max-relative entropy (24). It is clear that the operator $\hat{\rho}_i - \lambda_i\hat{\rho}$ in (29) becomes lower rank when

$$\lambda_i = 2^{-D_{\max}(\hat{\rho}||\hat{\rho}_i)}. \tag{30}$$

Max-relative entropy can also help us to formulate a closed-form criteria for the possibility of unambiguous state discrimination for the arbitrary ensemble $\{p_i, \hat{\rho}_i\}$: The state $\hat{\rho}_i$ can be unambiguously discriminated from the others if and only if

$$D_{\max}(\hat{\rho}_i||\hat{\rho}) = -\log_2 p_i. \tag{31}$$

Otherwise, $D_{\max}(\hat{\rho}_i||\hat{\rho}) < -\log_2 p_i$, which shows that the upper bound (25) is more rigorous than the Shannon entropy in (7).

The interpretation of (31) is simple: If $\hat{\rho} - p_i\hat{\rho}_i$ is lower rank, then $\hat{\rho}$ "loses" part of its support when $p_i\hat{\rho}_i$ is subtracted, thus projection on this part can be the POVM element which unambiguously identifies $\hat{\rho}_i$. When (31) holds for every i, the whole set can be unambiguously discriminated.

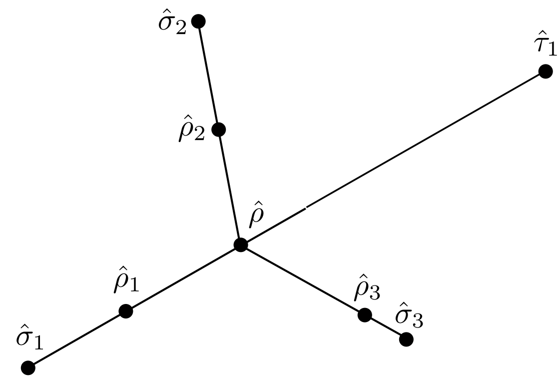

FIG. 2. Geometric representation of ensemble decomposition through extreme states. Subtracting $\hat{\rho}$ from $\hat{\rho}_i$ we get $\hat{\sigma}_i$. To construct a USD measurement for extreme states we move in the opposite direction obtaining $\hat{\tau}_i$. State $\hat{\sigma}_i$ is related to $D_{\max}(\hat{\rho}||\hat{\rho}_i)$, whereas state $\hat{\tau}_i$ is related to $D_{\max}(\hat{\rho}_i||\hat{\rho})$.

One can use (29) to make a decomposition of form (8) for any ensemble of states (see Fig. 2), but in a general case there may be no possibility of USD for the extreme states (a trivial case is when $\hat{\sigma}_i$ and $\hat{\rho}_i$ are the same). In order to do this, we can rewrite (29) as

$$\hat{\rho}_i = (1-\lambda_i)\hat{\sigma}_i + \lambda_i\hat{\rho}, \tag{32}$$

hence

$$\hat{\rho} = \sum_{i=1}^{N} p_i\hat{\rho}_i = \sum_{i=1}^{N} p_i(1-\lambda_i)\hat{\sigma}_i + \hat{\rho}\sum_{i=1}^{N} p_i\lambda_i, \tag{33}$$

and we obtain

$$\hat{\rho} = \frac{\sum_{i=1}^{N} p_i(1-\lambda_i)\hat{\sigma}_i}{1 - \sum_{i=1}^{N} p_i\lambda_i}. \tag{34}$$

Substituting (34) into (32) we get

$$\hat{\rho}_i = (1-\lambda_i)\hat{\sigma}_i + \frac{\lambda_i \sum_{k=1}^{N} p_k(1-\lambda_k)\hat{\sigma}_k}{1 - \sum_{k=1}^{N} p_k\lambda_k}. \tag{35}$$

Recall that here the values $\{\lambda_i\}$ are given by (30). The decomposition (35) has the form (8), but let us once again recall that the criteria for the possibility of USD may not hold for $\{\hat{\sigma}_i\}$ in the general case.

IV. COMPARISON WITH MAXIMUM CONFIDENCE MEASUREMENT

Let us compare our strategy with the maximum confidence measurement [16], because both measurement strategies can work when USD is not available, and, as we will show, sometimes our measurement outcomes do also provide maximum confidence besides providing mutual information gain.

The target functional to be maximized for maximum confidence measurement is the probability that the prepared state was $\hat{\rho}_i$, given that the outcome i was obtained

$$P(\hat{\rho}_i|i) = \frac{p_i \mathrm{Tr}(\hat{\rho}_i \hat{\Pi}_i)}{\mathrm{Tr}(\hat{\rho}\hat{\Pi}_i)}. \tag{36}$$

As shown in [16], the maximum confidence measurement can be viewed as a two-step process. We can rewrite (36) as follows:

$$P(\hat{\rho}_i|i) = \frac{p_i \mathrm{Tr}\big(\hat{A}_{\mathrm{succ}}\hat{\rho}_i\hat{A}^\dagger_{\mathrm{succ}}\hat{\Pi}'_i\big)}{\mathrm{Tr}\big(\hat{A}_{\mathrm{succ}}\hat{\rho}\hat{A}^\dagger_{\mathrm{succ}}\hat{\Pi}'_i\big)}, \tag{37}$$

where $\hat{\Pi}'_i = (\hat{A}^\dagger_{\mathrm{succ}})^{-1}\hat{\Pi}_i(\hat{A}_{\mathrm{succ}})^{-1}$ and

$$\hat{A}_{\mathrm{succ}} = \sqrt{\frac{P_{\mathrm{succ}}}{d}}\hat{\rho}^{-\frac{1}{2}}, \quad \hat{A}_{\mathrm{fail}} = \left(\hat{I} - \frac{P_{\mathrm{succ}}}{d}\hat{\rho}^{-1}\right)^{\frac{1}{2}} \tag{38}$$

are Kraus operators describing the first step of the measurement. In the case of a successful outcome any given input state $\hat{\rho}_i$ is transformed to

$$\hat{\rho}'_i = \frac{\hat{\rho}^{-\frac{1}{2}}\hat{\rho}_i\hat{\rho}^{-\frac{1}{2}}}{\mathrm{Tr}(\hat{\rho}_i\hat{\rho}^{-1})}. \tag{39}$$

Associated probability distribution is modified as

$$p'_i = \frac{P(\hat{\rho}_i)P(\mathrm{succ}|\hat{\rho}_i)}{P(\mathrm{succ})} = \frac{p_i}{d}\mathrm{Tr}(\hat{\rho}_i\hat{\rho}^{-1}). \tag{40}$$

Moreover,

$$\hat{\rho}' = \sum_{i=1}^{N} p'_i\hat{\rho}'_i = \frac{1}{d}\hat{I}. \tag{41}$$

Since the operator $\hat{\rho}'_i$ describing any given state commutes with that describing the other states in the ensemble $\hat{\rho}' - p'_i\hat{\rho}'_i$, taking into account $\hat{\rho}' \propto \hat{I}$, the same eigenvector corresponds to the largest eigenvalue of $\hat{\rho}'_i$ and the smallest eigenvalue of $\hat{\rho}' - p'_i\hat{\rho}'_i$. Hence, the maximum value of posterior probability (37) is attained, if $\hat{\Pi}'_i$ is a projector onto this eigenvector,

$$\max_{\hat{\Pi}_i} P(\hat{\rho}_i|i) = p_i\lambda_{\max}(\hat{\rho}^{-\frac{1}{2}}\hat{\rho}_i\hat{\rho}^{-\frac{1}{2}}), \tag{42}$$

where $\lambda_{\max}(\hat{\rho})$ is the maximum eigenvalue.

Using an alternative definition of max-relative entropy for two operators $\hat{\rho}_i$ and $\hat{\rho}$,

$$D_{\max}(\hat{\rho}_i||\hat{\rho}) = \log_2 \lambda_{\max}(\hat{\rho}^{-\frac{1}{2}}\hat{\rho}_i\hat{\rho}^{-\frac{1}{2}}), \tag{43}$$

the probability (42) can be expressed as

$$\max_{\hat{\Pi}_i} P(\hat{\rho}_i|i) = p_i 2^{D_{\max}(\hat{\rho}_i||\hat{\rho})}. \tag{44}$$

This relation shows an alternative operational meaning of the max-relative entropy: It quantifies the change of *a posteriori* state probability after the best possible measurement result was obtained. According to other operational interpretation of max-relative entropy (see [22,24]), the optimal Bayesian error probability is related to

$$H_{\min}(A|B)_{\rho_{AB}} = -\inf_{\sigma_B} D_{\max}(\rho_{AB}||I_A \otimes \sigma_B).$$

In our approach, the optimization is performed over all possible measurement operators, instead of POVMs, because with postselective measurement we are not restricted by the POVM condition $\sum_i \hat{M}_i = I$.

Now let us show that the same performance can be achieved with our geometric approach based on (29) and (35). If USD for the states $\{\hat{\sigma}_i\}$ is possible, it can be formulated as follows. We subtract given state $\hat{\rho}_i$ from $\hat{\rho}$ with some positive constant until we get a positive operator $\hat{\tau}_i$ with lower rank (see Fig. 2). As shown in Sec. III,

$$\hat{\tau}_i = \frac{\hat{\rho} - 2^{-D_{\max}(\hat{\rho}_i||\hat{\rho})}\hat{\rho}_i}{1 - 2^{-D_{\max}(\hat{\rho}_i||\hat{\rho})}}. \tag{45}$$

The corresponding POVM element maximizing posterior probability (36) is a projector onto the kernel of operator $\hat{\tau}_i$,

$$\hat{\Pi}_i = |\psi_i\rangle\langle\psi_i| \in \ker(\hat{\tau}_i). \tag{46}$$

Then

$$\max_{\hat{\Pi}_i} P(\hat{\rho}_i|i) = \frac{p_i\langle\psi_i|\hat{\rho}_i|\psi_i\rangle}{\langle\psi_i|\hat{\rho}|\psi_i\rangle} = p_i 2^{D_{\max}(\hat{\rho}_i||\hat{\rho})}. \tag{47}$$

Hence, we have shown that our geometrical approach based on (35) and Fig. 2 leads to the same results as ansatz (3) in [16].

If, in addition, the states $\{\hat{\sigma}_i\}$ in (35) allow for unambiguous state discrimination, both effects take place: Mutual information is above the Holevo quantity and all the conclusive measurement results provide maximum confidence. The only thing which is required for overcoming the Holevo bound is appropriate choice of success probabilities.

It may, however, happen that not all the outcomes for the measurement constructed in Theorem 1 provide maximum confidence. Let us describe this situation.

For the ensemble (8) we obtain

$$\hat{\rho} - \lambda\hat{\rho}_i = \sum_{k=1}^{K}\left(\sum_{j=1}^{N} p_j\alpha_{jk} - \lambda\alpha_{ik}\right)\hat{\sigma}_k, \tag{48}$$

and the max-relative entropy is equal to

$$D_{\max}(\hat{\rho}_i||\hat{\rho}) = -\log_2 \min_k \frac{\sum_{j=1}^{N} p_j\alpha_{jk}}{\alpha_{ik}}. \tag{49}$$

Thus, the maximum value of *a posteriori* probability (44) is

$$\max_{\hat{\Pi}_i} P(\hat{\rho}_i|i) = p_i \max_k \frac{\alpha_{ik}}{\sum_{j=1}^{N} p_j\alpha_{jk}}. \tag{50}$$

On the other hand, measurement (10) yields

$$P(\hat{\rho}_i|k) = \frac{p_i\alpha_{ik}}{\sum_{j=1}^{N} p_j\alpha_{jk}}. \tag{51}$$

Comparing (50) with (51) one can see that some of the measurement (10) outcomes yield maximal confidence, but some of them do not, and they may be considered as inconclusive, if the goal is maximization of (36). This is the difference between our approach and maximum confidence strategy, as well as the fact that the number K of conclusive measurement results in (8) may differ from the number N of states.

V. CONCLUSION

In this paper, we considered discrimination between quantum states, namely postselective measurements.

First, we have introduced the mutual information with postselection. We consider this value worthy of study, because it can quantify the quantum effect of increasing information after postselective measurement, which plays a key role in quantum cryptography, since communication over the public classical channel and discarding some positions is the key advantage of the legitimate users over the eavesdropper.

We have shown that this value can be greater than the Holevo quantity, and this effect also works when USD is impossible, including the case of two arbitrary noncommuting states. This result demonstrates that even individual measurements with postselection can be more effective than collective measurements without postselection. Thus the effect of postselective gain is demonstrated clearly. Remarkably, this effect can also take place when the number of outcomes is less than the number of states. Note that Theorem 1 states sufficient condition for overcoming the Holevo bound, but the necessary one is an open issue.

We also have shown the role of max-relative entropy for postselective measurements. Equation (44) states that it quantifies how confident one can be about the given state after obtaining the best possible result. Here we use the result of [16] which states the optimality for the proposed measurement, and the similarity of our approaches, while we used rather a geometric one. When criteria (31) holds, one can easily see from (44) that our conditional probability is 1, which corresponds to the unambiguous state discrimination case. Thus, (31) is a simple criteria for the possibility of USD. We have also shown that max-relative entropy plays the role of an upper bound for postselective mutual information; see (25).

Nevertheless, it should be noted that the ensemble (8) for which our result has been demonstrated, is rather specific. More general results for postselective information gain deserve further research, including the study of measurements over multiple copies of states (see, e.g., [25]).

We also find it interesting to generalize the proposed methods for arbitrary success probability, which has been done for USD [26].

ACKNOWLEDGMENTS

This work is supported by the Russian Science Foundation under Grant No. 20-71-10072.

[1] A. S. Holevo,Statistical decision theory for quantum systems, J. Multivariate Anal. **3**, 337 (1973).

[2] H. Yuen, R. Kennedy, and M. Lax, Optimum testing of multiple hypotheses in quantum detection theory, IEEE Trans. Inf. Theory **21**, 125 (1975).

[3] C. W. Helstrom, *Quantum Detection and Estimation Theory*, Vol. 84 (Academic Press, New York, 1976).

[4] S. M. Barnett and S. Croke, Quantum state discrimination, Adv. Opt. Photonics **1**, 238 (2009).

[5] J. A. Bergou, Discrimination of quantum states, J. Mod. Opt. **57**, 160 (2010).

[6] J. Bae and L.-C. Kwek, Quantum state discrimination and its applications, J. Phys. A: Math. Theor. **48**, 083001 (2015).

[7] A. S. Holevo, Bounds for the quantity of information transmitted by a quantum communication channel, Prob. Peredachi Inf. **9**, 3 (1973).

[8] I. Ivanovic, How to differentiate between non-orthogonal states, Phys. Lett. A **123**, 257 (1987).

[9] D. Dieks, Overlap and distinguishability of quantum states, Phys. Lett. A **126**, 303 (1988).

[10] A. Peres, How to differentiate between non-orthogonal states, Phys. Lett. A **128**, 19 (1988).

[11] J. A. Bergou, U. Futschik, and E. Feldman, Optimal Unambiguous Discrimination of Pure Quantum States, Phys. Rev. Lett. **108**, 250502 (2012).

[12] A. Peres and D. R. Terno, Optimal distinction between non-orthogonal quantum states, J. Phys. A: Math. Gen. **31**, 7105 (1998).

[13] G. Jaeger and A. Shimony, Optimal distinction between two non-orthogonal quantum states, Phys. Lett. A **197**, 83 (1995).

[14] A. Chefles, Unambiguous discrimination between linearly independent quantum states, Phys. Lett. A **239**, 339 (1998).

[15] Y. Feng, R. Duan, and M. Ying, Unambiguous discrimination between mixed quantum states, Phys. Rev. A **70**, 012308 (2004).

[16] S. Croke, E. Andersson, S. M. Barnett, C. R. Gilson, and J. Jeffers, Maximum Confidence Quantum Measurements, Phys. Rev. Lett. **96**, 070401 (2006).

[17] N. R. Kenbaev and D. A. Kronberg, Quantum measurement with post-selection for two mixed states, in *MIPT (PHYSTECH) - QUANT 2020*, edited by G. Lesovik, V. Vinokur, and M. Perelshtein, AIP Conf. Proc. No. 2362 (AIP, New York, 2021), p. 050001.

[18] A. Holevo, *Quantum Systems, Channels, Information: A Mathematical Introduction*, De Gruyter Studies in Mathematical Physics (De Gruyter, Berlin, 2012).

[19] M. Sasaki, K. Kato, M. Izutsu, and O. Hirota, Quantum channels showing superadditivity in classical capacity, Phys. Rev. A **58**, 146 (1998).

[20] M. Sasaki, S. M. Barnett, R. Jozsa, M. Osaki, and O. Hirota, Accessible information and optimal strategies for real symmetrical quantum sources, Phys. Rev. A **59**, 3325 (1999).

[21] H. Umegaki, Conditional expectation in an operator algebra. iv. entropy and information, Kodai Math. Sem. Rep. **14**, 59 (1962).

[22] N. Datta, Min- and max-relative entropies and a new entanglement monotone, IEEE Trans. Inf. Theory **55**, 2816 (2009).

[23] G. Gour, Uniqueness and optimality of dynamical extensions of divergences, PRX Quantum **2**, 010313 (2021).

[24] R. Konig, R. Renner, and C. Schaffner, The operational meaning of min- and max-entropy, IEEE Trans. Inf. Theory **55**, 4337 (2009).

[25] B. L. Higgins, A. C. Doherty, S. D. Bartlett, G. J. Pryde, and H. M. Wiseman, Multiple-copy state discrimination: Thinking globally, acting locally, Phys. Rev. A **83**, 052314 (2011).

[26] D. Kronberg, Increasing the distinguishability of quantum states with an arbitrary success probability, Proceedings of the Steklov Institute of Mathematics **313**, 113 (2021).