The optimal positive operator-valued measure for state discrimination

Wei Li¹,²,³ and Shengmei Zhao¹,⁴∗

¹Nanjing University of Posts and Telecommunications, Institute of Signal Processing and Transmission, Nanjing, 210003, China.
²Nanjing University of Posts and Telecommunications, Key Lab Broadband Wireless Communication and Sensor Network, Ministry of Education, Nanjing, 210003, China. and
³National Laboratory of Solid State Microstructures, Nanjing University, Nanjing 210093, China.

(Dated: September 27, 2021)

Abstract

Evaluating the amount of information obtained from non-orthogonal quantum states is an important topic in the field of quantum information. The commonly used evaluation method is Holevo bound, which only provides a loose upper bound for quantum measurement. In this paper, we provide a theoretical study of the positive operator-valued measure (POVM) for discriminating nonorthogonal states. We construct a generalized POVM measurement operation, and derive the optimal one for state discrimination by Lagrange multiplier method. With simulation, we find that the optimal POVM measurement provides a tight upper bound for state discrimination, which is significantly lower than that predicted by Holevo bound. The derivation of optimal POVM measurement will play an important role in the security research of quantum key distribution.

∗zhaosm@njupt.edu.cn
I. INTRODUCTION

Quantum measurement is a way to obtain information from unknown quantum systems, which is composed of a set of measurement operators that sum to the identity operator[1]. The simplest measurement operator is the projection operator, which is composed of a set of complete orthogonal bases, whose eigenvalue can be continuous or discrete, depending on the specific measured system. In quantum measurement, we are usually only interested in the state of the system, so we only pay attention to the probability of the result obtained by each measurement. A generalized quantum measurement is positive operator-valued measure (POVM)[2–5], in which the measurement operators can be non-orthogonal and completely described with the help of an ancillary state.

Quantum measurement has attracted extensive attentions in the field of quantum information. A typical example is evaluating the quantum channel attack in quantum key distribution (QKD) technology, where the eavesdropper carries out the channel attacks by interacting his ancillary states with the transmitted states[6, 7]. After the interaction, the eavesdroppers select appropriate quantum measurements for the stored ancillary quantum states to eavesdrop on as much information as possible about the transmitted quantum states. In quantum channel attacks, the ancillary quantum states stored by the eavesdropper are usually non-orthogonal. There are three main types of quantum channel attacks. The first is the coherent attack, which is a conceptual ideal high-dimensional channel attack scheme[8–12]. The key rate for this attack is provided by the entanglement purification protocol (EPP), which uses the quantum error correction code to purify the virtual Bell states, but EPP does not involve any quantum measurement. The second is the collective attack scheme, which is an asymptotically optimal quantum channel attack[13–18]. The eavesdropping capability of the eavesdropper is characterized by Holevo bound[19–21]. However, this method does not involve any specific measurement operation, so only an loose upper bound of the eavesdropping is provided. The third is the individual attack scheme[22–24], in which the eavesdropper’s attack capability is characterized by projection measurement. However, in individual attack, there is no detailed discussion on the generalized quantum measurement, so it is still uncertain whether the quantum measurement in this attack scheme is optimal. Therefore, finding the optimal quantum measurement scheme to characterize the tight upper bound of channel attack is an urgent problem to evaluate the capability of Eve’s
channel attacks.

In this paper, we give a theoretical study on the optimal POVM measurement for discriminating two non-orthogonal quantum states. Firstly, we construct a generalized POVM measurement, which is composed of bipartite interaction with an ancillary quantum state and followed a projection measurement of measured states after that interaction. Here, POVM measurement is equivalent to a classical channel, and its measurement capability is expressed as the mutual information. Next, we use Lagrange multiplier method to study the relationship between the dimension of the optimal projection measurement and the number of measured states, as well as find out the best POVM measurement. Finally, we compared the optimal POVM measurement with the Holevo bound for the quantum measurement of two non-orthogonal states.

II. GENERALIZED POVM MEASUREMENT FOR NON-ORTHOGONAL STATES

Imagine a game in which Alice tosses a coin, and send states $|p\rangle$ or $|q\rangle$ to Eve depending on whether the coin is head or tail. States $|p\rangle$ and $|q\rangle$ are non-orthogonal, and their inner product is $\langle p|q \rangle = \cos \alpha$, where $\alpha$ is the angle between states $|p\rangle$ and $|q\rangle$. Let’s further assume that the coin was tampered, the probability of occurrence of the head is $p$, the probability of occurrence of the tail is $q$, and they satisfy the normalization condition $p + q = 1$. So the mixed state that Eve received is

$$\rho = p\rho_p + q\rho_q, \quad (1)$$

where $\rho_p = |p\rangle \langle p|$ and $\rho_q = |q\rangle \langle q|$. Eve’s task is to find an appropriate POVM measurement, a generalized quantum measurement scheme, to measure the states $|p\rangle$ and $|q\rangle$, and infer whether Alice’s coin is head or tail according to the measurement results. So, what is the optimal POVM measurement to maximize the accuracy of Eve’s speculation?

A general POVM measurement can be formulated as a bipartite interaction with an ancillary state initiated as $|0\rangle_A$,

$$U |p\rangle |0\rangle_A = \sum_i^n M_i |p\rangle_i \langle i|_A = \sum_i^n \sqrt{p_i} |p_i\rangle \langle i|_A,$$

$$U |q\rangle |0\rangle_A = \sum_i^n M_i |q\rangle_i \langle i|_A = \sum_i^n \sqrt{q_i} |q_i\rangle \langle i|_A, \quad (2)$$

3
where $|i\rangle_A$ is a set of orthogonal bases with $\langle i|j\rangle_A = 0$, $M_i$ is a set of unitary operators that satisfies $\sum_i M_i^\dagger M_i = I$ with $I$ an identity operator, $p_i$ and $q_i$ are the probabilities of the occurrence of state $|i\rangle_A$ for measurement on $|p\rangle$ and $|q\rangle$ with $p_i = \langle p|M_i^\dagger M_i|p\rangle$, $q_i = \langle q|M_i^\dagger M_i|q\rangle$. Once state $|i\rangle_A$ is measured, states $|p_i\rangle$ and $|q_i\rangle$ should be distinguished to obtain the information about $|p\rangle$ and $|q\rangle$. This step is finished by a projection measurement by choosing a set of bases $|j\rangle$ with $j \in \{0, 1, 2, \cdots\}$. Then the total quantum measurement for states $|p\rangle$ and $|q\rangle$ can be formulated as a conditional probability matrix

$$p_M = \begin{bmatrix} p_1 |a_{1,0}|^2 & p_1 |a_{1,1}|^2 & \cdots & p_2 |a_{2,0}|^2 & p_2 |a_{2,1}|^2 & \cdots & p_n |a_{n,0}|^2 & p_n |a_{n,1}|^2 & \cdots \\ q_1 |b_{1,0}|^2 & q_1 |b_{1,1}|^2 & \cdots & q_2 |b_{2,0}|^2 & q_2 |b_{2,1}|^2 & \cdots & q_n |b_{n,0}|^2 & q_n |b_{n,1}|^2 & \cdots \end{bmatrix},$$

(3)

where $a_{i,j} = \langle j|p_i\rangle$ and $b_{i,j} = \langle j|q_i\rangle$, $\sum_j a_{i,j}^* b_{i,j} = \langle p_i|q_i\rangle = \cos \alpha_i$ with $\alpha_i$ the angle between states $|p_i\rangle$ and $|q_i\rangle$, the measurement matrix of the $k$-th sub-channel is

$$p_{M_k} = \begin{bmatrix} |a_{k,0}|^2 & |a_{k,1}|^2 & \cdots \\ |b_{k,0}|^2 & |b_{k,1}|^2 & \cdots \end{bmatrix}.$$  

(4)

As the phase factors in $|p\rangle$ and $|q\rangle$ have no effect on the measurement matrix and the module of the inner product between them, we can assume that both of their phases are 0, and the matrix elements $a_{i,j}$ and $b_{i,j}$ can be viewed as real numbers. Therefore, the POVM measurement matrix can be regarded as a channel, and the measurement scheme that can achieve the channel capacity is optimal.

III. DIMENSION OF OPTIMAL PROJECTION MEASUREMENT

Before deriving the channel capacity of the measurement matrix, we first study the relationship between the dimension of projection measurement and the number of measured states. We take the $i$-th subchannel as an example, where $\frac{pp_i}{pp_i + qq_i}$ and $\frac{qq_i}{pp_i + qq_i}$ are the normalized input probabilities of the channel, the mutual information for the $i$-th measurement matrix is

$$\chi_i = \sum_j \left( pp_i a_{i,j}^2 \log_2 \frac{(pp_i + qq_i) a_{i,j}^2}{pp_i a_{i,j}^2 + qq_i b_{i,j}^2} + qq_i b_{i,j}^2 \log_2 \frac{(pp_i + qq_i) b_{i,j}^2}{pp_i a_{i,j}^2 + qq_i b_{i,j}^2} \right),$$

(5)

which subject to conditions $\sum_j a_{i,j} b_{i,j} = \cos \alpha_i$ and $\sum_j \left( a_{i,j}^2 + b_{i,j}^2 \right) = 2$. The maximum value of $\chi_i$ can be calculated with Lagrange multiplier method. Here we construct the Lagrange
function for the $i$-th subchannel

$$L_i = \chi_i + \lambda_1 \left( \cos \alpha_i - \sum_j a_{i,j} b_{i,j} \right) + \lambda_2 \left( 2 - \sum_j a_{i,j}^2 + \sum_j b_{i,j}^2 \right), \quad (6)$$

where $\lambda_1$ and $\lambda_2$ are two constants. Around the maximum value of $\chi_i$, $L_i$ satisfies

$$\begin{align*}
\frac{\partial L_i}{\partial a_{i,j}} =& 2pp_i a_{i,j} \log_2 \left( \frac{pp_i + qq_i}{pp_i a_{i,j}^2 + qq_i b_{i,j}^2} \right) - \lambda_1 b_{i,j} - 2\lambda_2 a_{i,j}, \\
\frac{\partial L_i}{\partial b_{i,j}} =& 2qq_i b_{i,j} \log_2 \left( \frac{pp_i + qq_i}{pp_i a_{i,j}^2 + qq_i b_{i,j}^2} \right) - \lambda_1 a_{i,j} - 2\lambda_2 b_{i,j}.
\end{align*} \quad (7)$$

By summing the equations in Eq. (7), we have

$$2 \log_2 \left( \frac{pp_i a_{i,j}}{b_{i,j}} + \frac{qq_i b_{i,j}}{a_{i,j}} \right) + \frac{\lambda_1}{2pp_i q_i} \left( \frac{pp_i a_{i,j}}{b_{i,j}} + \frac{qq_i b_{i,j}}{a_{i,j}} \right) + \frac{\lambda_2 (pp_i + qq_i)}{pp_i q_i} + 2 \log_2 (pp_i + qq_i) = 0. \quad (8)$$

From Eq. (8), we can see that \( \frac{pp_i a_{i,j}}{b_{i,j}} + \frac{qq_i b_{i,j}}{a_{i,j}} \) is a function of \( pp_i qq_i \) and \( pp_i + qq_i \),

$$\frac{pp_i a_{i,j}}{b_{i,j}} + \frac{qq_i b_{i,j}}{a_{i,j}} = f (pp_i qq_i, pp_i + qq_i). \quad (9)$$

Once the values of \( pp_i \) and \( qq_i \) are provided, there are only two solutions for Eq. (9) that \( \frac{a_{i,j}}{b_{i,j}} = \{ g_1 (pp_i qq_i, pp_i + qq_i), g_2 (pp_i qq_i, pp_i + qq_i) \} \), whose value determines the placement of the projection bases with respect to \( |p_i\rangle \) and \( |q_i\rangle \).

Let’s set the plane formed by states \( |p_i\rangle \) and \( |q_i\rangle \) as \( \Gamma_i \), any projection basis \( |j\rangle \) can be decomposed into a component \( |j\rangle_\parallel \) parallel to \( \Gamma_i \) and a component \( |j\rangle_\perp \) perpendicular to \( \Gamma_i \), i.e. \( |j\rangle = c_j |j\rangle_\parallel + d_j |j\rangle_\perp \). Then we can get \( a_{i,j} = c_j \langle p_i |j\rangle_\parallel, b_{i,j} = c_j \langle q_i |j\rangle_\parallel \) and \( \frac{a_{i,j}}{b_{i,j}} = \frac{\langle p_i |j\rangle_\parallel}{\langle q_i |j\rangle_\parallel} \).

If we choose the first two bases \( |0\rangle \) and \( |1\rangle \) that satisfy \( \frac{\langle p_i |0\rangle_\parallel}{\langle q_i |0\rangle_\parallel} = g_1 (pp_i qq_i, pp_i + qq_i) \) and \( \frac{\langle p_i |1\rangle_\parallel}{\langle q_i |1\rangle_\parallel} = g_2 (pp_i qq_i, pp_i + qq_i) \), then a third basis \( |2\rangle \) that satisfies the same condition must parallel to \( |0\rangle_\parallel \) or \( |1\rangle_\parallel \) within \( \Gamma_i \), which will not guarantee the orthogonality of the bases. So a third basis does not exist, the dimension of optimal projection measurement is equal to the number of measured states.
IV. DERIVATION OF OPTIMAL POVM MEASUREMENT

With the conclusion of section III, the measurement matrix of Eq. (3) can be written as

\[
p_M = \begin{bmatrix}
p_1a_{1,0}^2 & p_1a_{1,1}^2 & p_2a_{2,0}^2 & p_2a_{2,1}^2 & \cdots & p_na_n^2 & p_n a_{n,1}^2 \\
q_1b_{1,0}^2 & q_1b_{1,1}^2 & q_2b_{2,0}^2 & q_2b_{2,1}^2 & \cdots & q_nb_n^2 & q_nb_{n,1}^2
\end{bmatrix},
\]

whose input probabilities are \( p \) and \( q \). Then the mutual information for the total POVM measurement matrix is

\[
\chi = \sum_{i,j} pp_i a_{i,j}^2 \log_2 \frac{ pp_i a_{i,j}^2 }{ pp_i a_{i,j}^2 + qq_i b_{i,j}^2 } + qq_i b_{i,j}^2 \log_2 \frac{ qq_i b_{i,j}^2 }{ pp_i a_{i,j}^2 + qq_i b_{i,j}^2 },
\]

which subject to conditions \( \sum_i \sqrt{ p_i q_i a_{i,j} b_{i,j} } = \cos \alpha \) and \( \sum_{i,j} ( pp_i a_{i,j}^2 + qq_i b_{i,j}^2 ) = 1 \), where \( i \in \{1, 2, \cdots, n\} \) and \( j \in \{0, 1\} \). Here again, we use Lagrange multiplier method to derive the maximum value of \( \chi \). The Lagrange function for \( \chi \) is

\[
L = \chi + \lambda_1 \left( \cos \alpha - \sum_i \sqrt{ p_i q_i a_{i,j} b_{i,j} } \right) + \lambda_2 \left( 1 - \sum_{i,j} ( pp_i a_{i,j}^2 + qq_i b_{i,j}^2 ) \right).
\]

Around the maximum value of \( \chi \), we have

\[
\frac{\partial L}{\partial a_{i,j}} = 2 pp_i a_{i,j} \log_2 \frac{ pp_i a_{i,j}^2 }{ pp_i a_{i,j}^2 + qq_i b_{i,j}^2 } - \lambda_1 p \sqrt{ p_i q_i b_{i,j} } - \lambda_2 2 pp_i a_{i,j},
\]

\[
\frac{\partial L}{\partial b_{i,j}} = 2 qq_i b_{i,j} \log_2 \frac{ pp_i a_{i,j}^2 }{ pp_i a_{i,j}^2 + qq_i b_{i,j}^2 } - \lambda_1 p \sqrt{ p_i q_i a_{i,j} } - \lambda_2 2 qq_i b_{i,j}.
\]

By summing the equations in Eq. (13), we get

\[
\log_2 \left( p \sqrt{ \frac{ p_i a_{i,j} }{ q_i b_{i,j} } } + q \sqrt{ \frac{ q_i b_{i,j} }{ p_i a_{i,j} } } \right) + \frac{\lambda_1}{4} \left( p \sqrt{ \frac{ p_i a_{i,j} }{ q_i b_{i,j} } } + q \sqrt{ \frac{ q_i b_{i,j} }{ p_i a_{i,j} } } \right) + \lambda_2 = 0.
\]

From Eq. (14), we can see that \( p \sqrt{ \frac{ p_i a_{i,j} }{ q_i b_{i,j} } } + q \sqrt{ \frac{ q_i b_{i,j} }{ p_i a_{i,j} } } \) is a constant. Let’s set \( p \sqrt{ \frac{ p_i a_{i,j} }{ q_i b_{i,j} } } + q \sqrt{ \frac{ q_i b_{i,j} }{ p_i a_{i,j} } } = \gamma \), we have the following equality

\[
\frac{ pp_i a_{i,j}^2 + qq_i b_{i,j}^2 }{ b_{i,j} } = \frac{ pp_i a_{i,j}^2 + qq_i b_{i,j}^2 }{ a_{i,j} b_{i,j} } = \gamma \sqrt{ p_i q_i }.
\]

here \( p \) and \( q \) are regarded as known constants, \( p_i \) and \( q_i \) are variables that depend on the subchannel index \( i \). From Eq. (15), we further get

\[
\sum_j ( pp_i a_{i,j}^2 + qq_i b_{i,j}^2 ) = pp_i + qq_i = \sum_j \gamma \sqrt{ p_i q_i a_{i,j} b_{i,j} } = \gamma \sqrt{ p_i q_i } \cos \alpha_i.
\]
Because the total mutual information of the POVM measurement is the sum of the mutual information of each subchannel, and the subchannels are independent of each other, when $\chi$ reaches its maximum, $\chi_i$ for the $i$-th subchannel should reach its maximum as well. Combine Eqs. (9) and (15), we have $pp_i + qq_i = h\left(\sqrt{p_iq_i}\right)$. It’s easy to see that polynomials $pp_i + qq_i$ and $\sqrt{p_iq_i}$ are in the same order, then $pp_i + qq_i \propto \sqrt{p_iq_i}$, so we can conclude that $p\sqrt{\frac{p_i}{q_i}} + q\sqrt{\frac{q_i}{p_i}}$ is a constant,

$$p\sqrt{\frac{p_i}{q_i}} + q\sqrt{\frac{q_i}{p_i}} = \gamma_1 = \gamma \cos \alpha_i. \quad (17)$$

It is easy to see that the value of $\cos \alpha_i$ is independent of the subchannel index $i$. Here we set $\cos \alpha_i = \gamma_2$, then the equality $\sum_i \sqrt{p_iq_i} \cos \alpha_i = \gamma_2 \sum_i \sqrt{p_iq_i} = \cos \alpha$ can be obtained. Because $\sum_i \sqrt{p_iq_i} \leq 1$, so $\gamma_2 = \frac{\cos \alpha}{\sum_i \sqrt{p_iq_i}} \geq \cos \alpha$, where the equality holds when $p_i = q_i$. As the discrimination of states $|p_i\rangle$ and $|q_i\rangle$ depends on the angle between them, the smaller $\gamma_2$ is, the larger $\alpha_i$ we have, the more we can distinguish $|p_i\rangle$ and $|q_i\rangle$. So the optimal POVM measurement is that for each subchannel, $p_i = q_i$, $\langle p_i|q_i\rangle = \cos \alpha_i = \cos \alpha$ and $p_{a_{i,j}} + q_{b_{i,j}}$ is independent of the subchannel index $i$. In this case, the optimal POVM measurement is equivalent to a projection measurement on the initial states. Therefore, in this game, Eve’s optimal POVM measurement strategy is to use the projection measurement that meets the above conditions, which could maximize the accuracy of her guess about the result of Alice’s coin toss.

Another more practical role played by POVM measurement is to evaluate the eavesdropping capability of quantum channel attacks by distinguishing non-orthogonal states in the security analysis of QKD. Next, we compare POVM measurement and Holevo bound, which are two commonly used information quantization methods in the field of quantum information. For two equiprobability mixed states $|p\rangle$ and $|q\rangle$, the maximum amount of information that can be obtained by POVM measurement is $\chi_P = 1 - H\left(\frac{1 + \sin \alpha}{2}\right)$, while the maximum amount of information that can be obtained by Holevo bound is $\chi_H = H\left(\frac{1 + \cos \alpha}{2}\right)$, where $H(x)$ is Shannon entropy with $H(x) = -x \log_2 x - (1 - x) \log_2 (1 - x)$ and $\langle p|q\rangle = \cos \alpha$. Fig.1 shows the comparison of $\chi_P$ and $\chi_H$ with respect to $\alpha$, where Holevo bound is believed to provide an upper bound for information extraction\[19\]. The black solid line is the amount of information evaluated by the optimal POVM measurement the red solid line is that evaluated by Holevo bound, and the blue solid line is the difference between
FIG. 1. Comparison of the amount of information evaluated by optimal POVM measurement $\chi_P$ and Holevo bound $\chi_H$ as well as the difference between them $\chi_H - \chi_P$ from two non-orthogonal states with respect to the angle $\alpha$ between them.

From this figure, we can see that the information evaluated by Holevo bound is always larger than that evaluated by POVM measurement except for $\alpha = 0$ or $\alpha = \pi/2$, where these two states can not be distinguished or can be completely distinguished. Therefore, the Holevo bound provides a loose upper bound of eavesdropping capability of quantum channel attacks, while the optimal POVM measurement provides a tight bound of that. With the optimal POVM measurement derived above, a higher key rate can be obtained compared with the methods based on EPP and Holevo bound.

V. CONCLUSION

In this paper, We derive the optimal POVM measurement scheme for two arbitrary non-orthogonal states. Here we find that the measurement matrix of POVM measurement can be equivalent to a classical channel, projection measurements one of the optimal POVM measurement, and the dimension of projection measurement is equal to the number of states to be measured. Compared with Holevo bound, the optimal POVM measurement provides a tight bound on the amount of information obtained from non-orthogonal states. Based on this, the optimal POVM measurement will play an important role in improving the key rate of practical QKD.
ACKNOWLEDGMENTS

This work is supported by China Postdoctoral special funding project (2020T130289), the National Natural Science Foundation of China (No. 61871234).

[1] M. A. Nielsen and I. Chuang, “Quantum computation and quantum information,” (2002).
[2] H. E. Brandt, American Journal of Physics 67, 434 (1999).
[3] S. Hamieh, R. Kobes, and H. Zaraket, Physical Review A 70, 052325 (2004).
[4] A. Winter, Communications in mathematical physics 244, 157 (2004).
[5] M. Ziman, Physical Review A 77, 062112 (2008).
[6] V. Scarani, H. Bechmann-Pasquinucci, N. J. Cerf, M. Dušek, N. Lütkenhaus, and M. Peev, Reviews of modern physics 81, 1301 (2009).
[7] F. Xu, X. Ma, Q. Zhang, H.-K. Lo, and J.-W. Pan, Reviews of Modern Physics 92, 025002 (2020).
[8] H.-K. Lo and H. F. Chau, science 283, 2050 (1999).
[9] P. W. Shor and J. Preskill, Physical review letters 85, 441 (2000).
[10] L. Sheridan, T. P. Le, and V. Scarani, New Journal of Physics 12, 123019 (2010).
[11] H.-K. Lo, M. Curty, and B. Qi, Physical review letters 108, 130503 (2012).
[12] X. Ma, P. Zeng, and H. Zhou, Physical Review X 8, 031043 (2018).
[13] E. Biham and T. Mor, Physical Review Letters 78, 2256 (1997).
[14] E. Biham, M. Boyer, G. Brassard, J. Van De Graaf, and T. Mor, Algorithmica 34, 372 (2002).
[15] M. Boyer, R. Gelles, and T. Mor, Algorithms 2, 790 (2009).
[16] A. Acín, N. Brunner, N. Gisin, S. Massar, S. Pironio, and V. Scarani, Physical Review Letters 98, 230501 (2007).
[17] S. Pironio, A. Acín, N. Brunner, N. Gisin, S. Massar, and V. Scarani, New Journal of Physics 11, 045021 (2009).
[18] W. Li, L. Wang, and S. Zhao, Scientific reports 9, 1 (2019).
[19] A. S. Holevo, Problemy Peredachi Informatsii 9, 3 (1973).
[20] M. E. Shirokov, Problems of Information Transmission 55, 201 (2019).
[21] W. Roga, M. Fannes, and K. Życzkowski, Physical review letters 105, 040505 (2010).
[22] C. A. Fuchs, N. Gisin, R. B. Griffiths, C.-S. Niu, and A. Peres, Physical Review A 56, 1163 (1997).

[23] R. B. Griffiths and C.-S. Niu, Physical Review A 56, 1173 (1997).

[24] H. Bechmann-Pasquinucci and N. Gisin, Physical Review A 59, 4238 (1999).